

Université de Sherbrooke

Identification des résidus contribuant à
l'interaction hnRNP A1- hnRNP A1

Par Benoît Labrecque
Département de microbiologie et d'infectiologie
Laboratoire du docteur Benoit Chabot

Mémoire présenté à la Faculté de Médecine
En vue de l'obtention d'un grade de
Maître ès sciences (M.Sc.)

16 juillet 2003



National Library
of Canada

Bibliothèque nationale
du Canada

Acquisitions and
Bibliographic Services

Acquisitions et
services bibliographiques

395 Wellington Street
Ottawa ON K1A 0N4
Canada

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

ISBN: 0-612-90603-5

Our file Notre référence

ISBN: 0-612-90603-5

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this dissertation.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de ce manuscrit.

While these forms may be included in the document page count, their removal does not represent any loss of content from the dissertation.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

RESUME

HnRNP A1 est une protéine impliquée dans la sélection des sites 5' d'épissage *in vitro* et *in vivo*. Il est connu que son domaine riche en glycines est important pour médier son activité dans l'épissage alternatif. Le modèle « looping-out » a été proposé comme mécanisme d'action de A1; lorsque A1 lie des sites de haute affinité localisés dans la région intronique de part et d'autre d'un exon alternatif, l'exclusion de celui-ci est favorisé par une interaction A1-A1 impliquant le domaine riche en glycines. Les régions et les acides aminés impliqués dans l'interaction A1-A1 restent méconnus. Une mutagenèse aléatoire de l'ADNc de A1, suivie d'une approche génétique chez la bactérie basée sur la répression traductionnelle, nous a permis d'aborder l'importance du domaine riche en glycines et du domaine RRM2 dans la liaison coopérative de la protéine hnRNP A1 à l'ARN. La majorité des mutants identifiés montrent un défaut au niveau de l'interaction A1-A1 lorsque quantifiés dans un essai double-hybride chez la levure. Plusieurs régions dans les domaines riches en glycines et dans le domaine RRM2 pourraient ainsi participer à l'interaction A1-A1. L'analyse sur gel de retardement montre que certaines mutations dans le domaine RRM2 atténuent grandement la capacité de A1 à lier les acides nucléiques, suggérant que le RRM1 seul est insuffisant pour une liaison forte de A1. Enfin, l'étude de ces mutants dans la sélection de sites d'épissage 5' *in vitro* suggère que l'interaction A1-A1 est impliquée dans le mécanisme de modulation de l'épissage alternatif. Donc, les acides aminés participants à l'interaction A1-A1 influencent hnRNP A1 dans son activité sur l'épissage alternatif.

TABLE DES MATIERES

Table des matières.....	I
Liste des figures.....	III
Liste d'abréviations	IV
Résumé.....	
Introduction.....	1
<i>Épissage</i>	<i>1</i>
<i>L'épissage constitutif.....</i>	<i>3</i>
<i>Épissage alternatif.....</i>	<i>8</i>
<i>Les modulateurs de l'épissage alternatif.....</i>	<i>9</i>
<i>Protéines SR.....</i>	<i>10</i>
<i>Protéines hnRNPs.....</i>	<i>11</i>
<i>HnRNP A1.....</i>	<i>12</i>
<i>RRMs.....</i>	<i>13</i>
<i>Le domaine riche en glycines.....</i>	<i>17</i>
<i>HnRNP A1 et l'épissage alternatif.....</i>	<i>18</i>
Matériel et méthodes.....	26
<i>Constructions plasmidiques.....</i>	<i>26</i>
<i>Mutagenèse par PCR.....</i>	<i>27</i>
<i>Répression traductionnelle</i>	<i>28</i>
<i>Transcrits radiomarqués pour réaction d'épissage in vitro et gel de retardement</i>	<i>28</i>
<i>Marquage d'oligonucléotides pour gel de retardement</i>	<i>29</i>
<i>Déplétion d'extraits nucléaires en hnRNP A1</i>	<i>29</i>
<i>Détection de hnRNP A1 par immunobuvardage (Western blot).....</i>	<i>29</i>
<i>Production et purification de protéines recombinantes.....</i>	<i>30</i>
<i>Gel de retardement</i>	<i>32</i>
<i>Réaction d'épissage in vitro.....</i>	<i>32</i>
<i>Technique double-hybride</i>	<i>33</i>
Résultats.....	34
<i>Détection de l'interaction A1-A1 par répression traductionnelle</i>	<i>35</i>

<i>Séquences</i>	39
<i>La liaison aux acides nucléiques est affectée par certaines mutations</i>	41
<i>Interaction A1-A1 des mutants selon le double hybride chez la levure</i>	45
<i>Certaines mutations dans le domaine riche en glycines affectent la sélection des sites d'épissage 5'</i>	48
Discussion	53
<i>Validité du système de répression traductionnelle</i>	53
<i>L'interaction A1-A1 nécessite une liaison à une séquence d'acide nucléique simple brin</i>	54
<i>Le rôle du RRM2</i>	55
<i>Le domaine riche en glycines influence l'interaction protéine-protéine</i>	57
<i>Rôle du domaine riche en glycines dans l'épissage</i>	59
<i>HnRNP A1 modélisée dans l'épissage alternatif</i>	60
Remerciements	62
Bibliographie	63
Annexe I	74
Annexe II	83

LISTE DES FIGURES

Figure 1	2
Figure 2	4
Figure 3	6
Figure 4	15
Figure 5	21
Figure 6	23
Figure 7	37
Figure 8	40
Figure 9	42
Figure 10	44
Figure 11	47
Figure 12	50
Figure 13	51
Figure 14	52

LISTE D'ABRÉVIATIONS

a.a.	acide aminé
ADN	acide désoxyribonucléique
ADNc	acide désoxyribonucléique complémentaire
ARN	acide ribonucléique
BSA	“bovine serum albumin”
°C	degré Celsius
CTD	“Carboxy-terminal domain”
DTT	dithiothréitol
EDTA	Éthylène diamide tétraacétique acide
ESE	“Exonic Splicing Enhancer”
ESS	“Exonic Splicing Silencer”
GST	glutathione-S-transférase
His	histidine
hnRNP	“heterogenous nuclear ribonucleoprotein”
IPTG	isopropyl-1-thio- β -D-galactosidase
ISE	“Intronic Splicing Enhancer”
ISS	“Intronic Splicing Silencer”
kDa	kiloDalton
M	molaire (mol/litre)
min	minute
mg	milligramme (10^{-3} g)

ml	millilitre (10^{-3} l)
nm	nanomètre (10^{-9} m)
pmole	picomole (10^{-12} mole)
PMSF	3-phénylméthylsulfonyl fluorure
PTB	“polyrimidine-tract binding protein”
RBD	“RNA binding domain”
rpm	rotation par minute
RRM	“RNA Recognition Motif”
sec	seconde
SELEX	“Selective Evolution of Ligand by Exponential enrichment”
SF1	“Splicing factor 1”
snRNP	“small nuclear ribonucleoprotein”
Tris	tris (hydroxyl méthy) aminoéthane
U2AF	“U2 auxiliary factor”
µg	microgramme (10^{-6} g)
µl	microlitre (10^{-6} l)
µM	micromolaire (10^{-6} M)
UP1	“unwinding protein 1”
wt	“wild type”, de type sauvage

INTRODUCTION

Épissage

La cellule constitue la plus petite unité biologique pouvant vivre de manière indépendante. À travers l'évolution, elle est devenue structurellement plus complexe en fonction de ses besoins ainsi que par ses interactions avec son environnement. Afin de survivre, la cellule nécessite une fine régulation de l'expression de ses gènes, et cette régulation apparaît à divers niveaux chez les eucaryotes. La régulation s'étend de la transcription des gènes à la maturation des ARNs pré-messagers en passant au transport de ceux-ci et à leur dégradation. Parmi la panoplie de processus de maturation, l'épissage consiste l'un des mécanismes le plus étudié. L'épissage se définit par l'enlèvement des introns d'un ARN pré-messager, suivi de la réunion des exons, régions encodant la future protéine. L'ARNm ainsi formé sera transporté hors du noyau vers le cytoplasme et sera traduit par les ribosomes pour fabriquer la protéine voulue. L'épissage alternatif est une variante de l'épissage constitutif (Figure 1). À partir d'un même ARN pré-messager, il permet la maturation de différents ARNm menant à des protéines possédant souvent des fonctions distinctes.

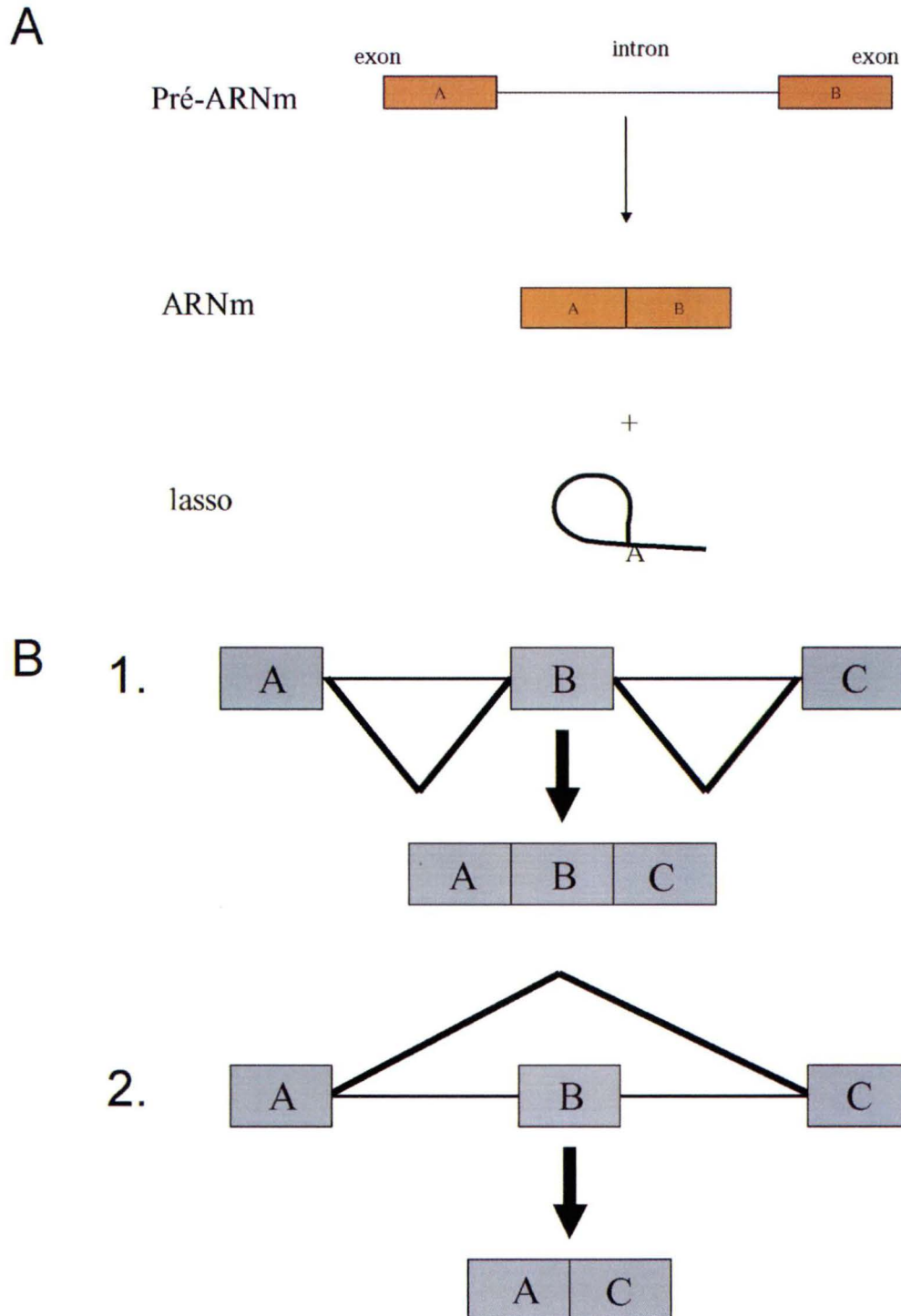


Figure 1. A. **Épissage constitutif.** Les introns sont enlevés et les exons unis pour former l'ARNm. Les introns, sous forme de lasso, sont menés vers les voies de dégradation. B. **Épissage alternatif.** Ce mécanisme procède à l'inclusion ou l'exclusion d'un exon alternatif, menant à des ARNm différents et donc des protéines différentes. L'utilisation des sites d'épissage internes permet l'inclusion de l'exon alternatif (1), alors que l'utilisation exclusive des sites d'épissage distaux permet l'exclusion de l'exon alternatif (2).

La découverte de l'épissage et de l'épissage alternatif a été faite grâce à des études des ARNm tardifs de l'adénovirus (Berget et al., 1977; Chow et al., 1977). Depuis, le génôme humain a été séquencé et le potentiel de l'épissage alternatif à générer la diversification des protéines constitue la prochaine étape à cartographier puisque les 32000 gènes dénombrés ne suffisent pas à spécifier l'ensemble des protéines. Le gène Ds-CAM chez la drosophile représente le cas le plus éloquent du pouvoir de diversité issu de l'épissage alternatif. Pour ce seul gène, on évalue le nombre d'isoformes possibles à 38000 (Schmucker et al., 2000). Chez les mammifères, 99% des gènes contiennent des introns, et il est maintenant estimé qu'environ 60% des gènes humains subissent au moins un événement d'épissage alternatif (McPherson et al., 2001). Ces statistiques alimentent les études dans la compréhension des mécanismes de l'épissage, surtout si l'on tient compte qu'environ 15% des maladies génétiques humaines sont causées par des mutations ponctuelles troublant des signaux d'épissage (Caceres et Kornblihtt, 2002).

L'épissage constitutif

Le mécanisme de l'épissage requiert un ensemble de facteurs bien déterminé avant de s'exécuter. D'abord, les sites d'épissage, les jonctions entre l'exon et l'intron, doivent être identifiées. Chez les mammifères, les sites d'épissage 5' correspondent à la séquence consensus GURAGU et les sites d'épissage 3' à la séquence YAG. De 18 à 40 nucléotides en amont du site d'épissage 3', on retrouve le site de branchement qui se conforme à la séquence consensus YNYURAC, (Y = pyrimidine, R = purine et N = un nucléotide), et est habituellement suivie d'une séquence riche en pyrimidines

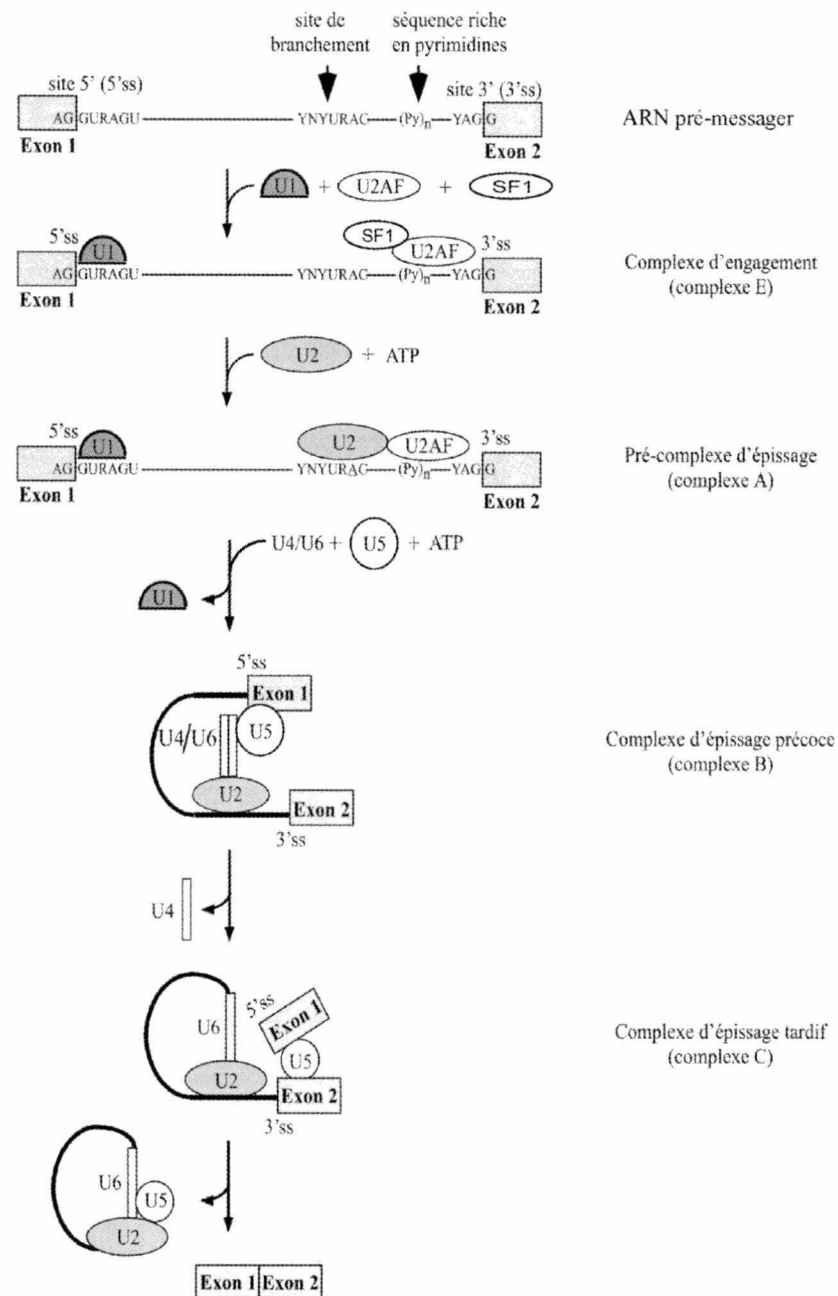


Figure 2. Formation du spliceosome. L'épissage est médié par la formation du spliceosome suite à l'addition de 5 snRNPs (U1, U2, U4, U5 et U6) sur l'ARN prémessager. Le spliceosome permet de catalyser 2 réactions de transestérification. La réaction d'épissage terminée, les snRNPs se recyclent dans un autre spliceosome.

(Figure 2, ARN pré-messager) (Smith et Valcárcel, 2000). Malgré le consensus, les séquences sont dégénérées chez les eucaryotes supérieurs donnant accès à des sites d'épissage d'affinité variée. Biochimiquement, la réaction d'épissage se traduit par deux réactions de transestérification (Brow et al, 2002). Lors de la première étape, une attaque nucléophile du groupement hydroxyl en 2' de l'adénosine du site de branchement au groupement phosphate du site d'épissage 5' détruit la liaison phosphodiester de la jonction exon-intron. L'intron prend la forme d'un lasso grâce à la nouvelle liaison phosphodiester 2'-5'. Ensuite, le groupement 3' hydroxyl de l'exon en 5' procède à une seconde attaque nucléophile sur le phosphate de la jonction exon-intron en 3'. Cette dernière transestérification permet de lier les deux exons et libère l'intron sous forme de lasso, qui sera ultérieurement dégradé (Figure 3).

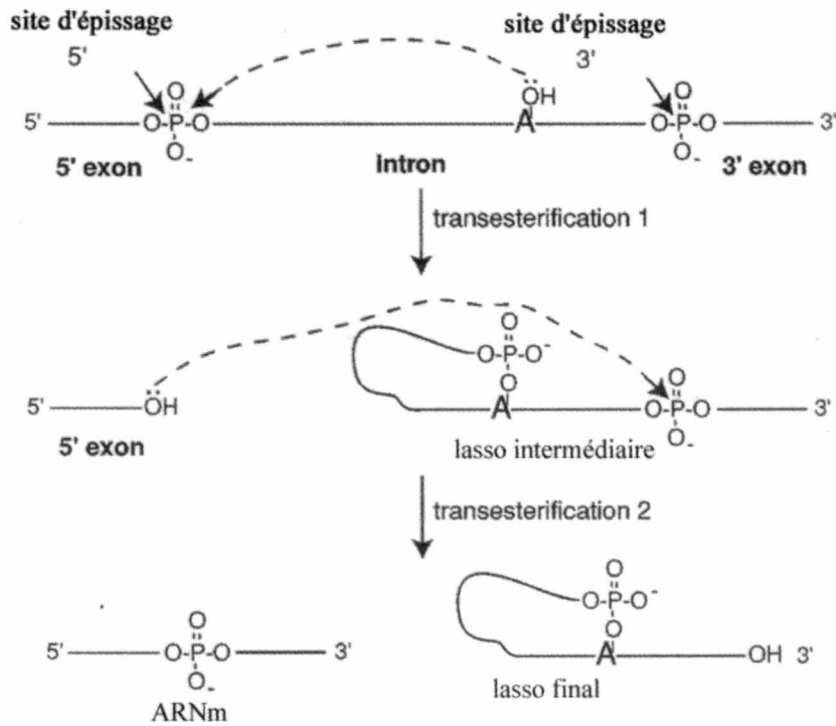


Figure 3. **Processus chimique de la réaction d'épissage.** (tiré de Brow, 2002)

La réaction d'épissage est catalysée par un complexe multi-ribonucléoprotéique de 60S appelé spliceosome. Le spliceosome est constitué de 5 snRNPs et d'un ensemble de protéines accessoires (Krämer, 1996; Will et Lührmann, 1997). L'assemblage du spliceosome débute par la liaison du U1 snRNP à la séquence consensus du site d'épissage 5' par appariement ARN-ARN (**Figure 2**). À l'autre extrémité, U2AF, un facteur d'épissage composé des sous-unités U2AF³⁵ et U2AF⁶⁵, se lie à la séquence riche en pyrimidine. Le facteur d'épissage SF1 reconnaît le site de branchement par son domaine KH avec l'aide de U2AF (Peled-Zehavi et al., 2001). Les protéines SR participent également à cet assemblage du complexe d'engagement (complexe E). La liaison de U2AF au site d'épissage 3' favorise la liaison du snRNP U1 au prochain site d'épissage 5' en aval. Ceci permet de cibler l'exon avant son excision par le spliceosome et ce mécanisme est connu sous le nom « exon definition » (Wang et al., 1995). Par la suite, U2AF⁶⁵ contribue au déplacement de SF1 par la liaison de U2 au site de branchement via une interaction ARN-ARN, pour former le complexe du pré-spliceosome (complexe A). Cette interaction a la particularité d'exposer davantage l'adénosine du site de branchement participant à la première attaque nucléophile. Le tri-snRNP U4/U5/U6 intervient ensuite pour former le complexe d'épissage précoce (complexe B). Ce complexe subit quelques réarrangements pour devenir un complexe actif pour la catalyse des deux réactions de transestérification (complexe C). U6 abandonne son interaction ARN-ARN avec U4 pour former une liaison avec U2, et il déplace aussi U1 pour interagir avec le site d'épissage 5'. Malgré cette description, le modèle sur l'assemblage du spliceosome n'est pas encore établi de

façon définitive (Brow, 2002). L'existence d'un spliceosome pré-assemblé a maintenant été rapportée (Stevens et al., 2002). Aussi, il existe d'autres classes mineures de spliceosomes (Otake et al., 2002).

Épissage alternatif

L'épissage alternatif peut se définir par l'utilisation différentielle d'exons avec comme conséquence que l'on peut produire différentes isoformes d'une protéine à partir d'un même type d'ARN pré-messager. Les isoformes d'une protéine peuvent être produites spécifiquement selon le type cellulaire, le tissu, le sexe et le stade développemental. Le principe de l'épissage alternatif repose sur la sélectivité des sites d'épissage de l'ARN précurseur. La dégénérescence des sites d'épissage chez les eucaryotes supérieurs attribue une affinité relative à chaque site. Donc, un site d'épissage, se conformant de près à la séquence consensus, sera un site choisi plus facilement par la machinerie d'épissage qu'un site ayant une faible similarité avec la séquence consensus. Ces sites forts sont souvent associés aux exons épissés constitutivement. Les sites plus faibles avoisinant majoritairement les exons alternatifs ont davantage de difficultés à recruter les facteurs d'épissage. L'utilisation d'un site d'épissage faible peut être augmentée ou diminuée par certains facteurs d'épissage, alors qu'un site d'épissage fort est régulé plus difficilement.

Les modulateurs de l'épissage alternatif

La modulation de l'épissage alternatif implique une multitude d'éléments *cis* et de facteurs *trans*. À prime abord, la nature de l'ARN pré-messager est importante selon l'affinité de ses sites d'épissage, la taille des exons, la localisation et le nombre de sites de branchement pour un événement d'épissage. L'identité du promoteur agit également sur la sélection d'un site d'épissage, soit par la processivité de l'ARN polymérase II, soit par le recrutement des protéines SR (Maniatis et Reed, 2002). La présence de codons stop influencerait possiblement le choix du site 5' afin d'utiliser le bon cadre de lecture déterminé lors de l'expression du gène (Li et al., 2002). On retrouve également des séquences activatrices exoniques (ESE) ou introniques (ISE), et des séquences inhibitrices exoniques (ESS) ou introniques (ISS). Ces séquences stimulent ou inhibent l'utilisation d'un site d'épissage en recrutant et stabilisant des facteurs. Les protéines SR sont un exemple d'activateurs lorsqu'elles lient un ESE afin de favoriser l'épissage à un site adjacent (Lavigneur et al., 1993). La formation de structures secondaires a déjà été rapportée comme élément intronique pouvant contrer l'utilisation d'un site d'épissage. Par exemple, la formation d'un duplex ARN-ARN dans la région intronique de l'ARN pré-messager de hnRNP A1, permet de réguler le site d'épissage 5' de l'exon 7B (Blanchette et Chabot, 1997). Les deux modulateurs *trans* les plus communs à l'épissage alternatif sont les protéines SR et les protéines hnRNPs. Leur concentration dans le noyau et leur affinité relative pour des séquences spécifiques pour l'ARN dirigent l'inclusion ou l'exclusion d'un exon alternatif (Chabot, 1996). Enfin, des agents chimiques, tels des solvants organiques

comme le diméthyl sulfoxyde et le diméthyl formamide ont la propriété d'influencer la sélection des sites d'épissage 5' (Bolduc et al., 2001). Le mécanisme exact reste inconnu, mais l'activation des protéines SR par une phosphorylation accrue en présence de ces solvants est suspectée.

Protéines SR

Les protéines SR forment une famille caractérisée par la présence d'un domaine riche en sérine et arginine (domaine RS) en position carboxy-terminale. En position N-terminale, elles peuvent contenir un ou deux domaines de liaison à l'ARN de type RRM « RNA Recognition Motif ». Leur rôle est redondant au sein de l'épissage constitutif. Puisque les protéines SR peuvent interagir avec le domaine CTD « Carboxy-Terminal Domain » de l'ARN polymérase II, les protéines SR permettraient de coupler l'épissage à la transcription (Hirose et Manley, 2000). Les différents membres de la famille des protéines SR ont généralement la capacité de moduler l'épissage alternatif en liant spécifiquement les ESE par leur domaine de liaison à l'ARN. Le recrutement des SR au ESE permet de favoriser l'utilisation d'un site d'épissage 3' ou 5' en recrutant les facteurs d'épissage essentiels (Graveley, 2000). Cependant, une protéine SR peut également avoir une activité inhibitrice, telle SRp30c qui lie un élément intronique dans l'ARN pré-messager de A1; il y a alors répression de l'utilisation du site 3' situé en aval de cet élément (Simard et Chabot, 2002). Le niveau d'expression de chaque membre de la famille des protéines SR varie en fonction du type cellulaire, suggérant une modulation de l'expression des gènes

spécifiques au tissu (Zahler et al., 1992). L'activité des protéines SR est liée étroitement à leur degré de phosphorylation (Misteli, 1999).

Protéines hnRNPs

Contrairement aux protéines SR, les protéines hnRNPs ont souvent une activité inhibitrice au niveau de l'épissage alternatif. Les protéines hnRNPs « heterogenous nuclear ribonucleoparticles » sont parmi les protéines nucléaires les plus abondantes dans les cellules eucaryotes. Elles sont caractérisées par leur capacité à lier spécifiquement certaines séquences d'ARN selon différentes affinités par des domaines RRM (ou RBD) (Chabot et al., 2003). Jusqu'à maintenant, on dénombre environ une trentaine de protéines hnRNPs dans les cellules de vertébrés dont les tailles varient entre 34 kDa et 120 kDa. Elles participent à diverses activités de maturation de l'ARN. On les retrouve associées aux transcrits récemment polymérisés par l'ARN polymérase II, sous forme de complexes (Choi et Dreyfuss, 1984). On retrouve les protéines hnRNPs à diverses étapes de l'expression des gènes : la transcription, la polyadénylation, l'épissage, le transport et la stabilité des acides nucléiques. HnRNP A1, F, H et I sont les molécules impliquées dans l'épissage alternatif les mieux étudiées. HnRNP I, plus connue sous le nom de PTB, est caractérisée par 4 domaines RRM reconnaissant spécifiquement des séquences riches en uridine et cytosine habituellement situées près du site 3' d'épissage. La plupart du temps, il a été observé que hnRNP I régule négativement l'épissage d'un exon (Black, 2003). Il existe au moins un cas où hnRNP I est liée à une activité positive dans

l'épissage (Lou et al., 1999). La protéine hnRNP H est fortement homologue à hnRNP F. Chacune contient 3 domaines RRM, reconnaissant une séquence d'ARN GGGA (Caputi et Zahler, 2001), et 2 domaines riches en glycines. HnRNP H et F peuvent être impliquées dans l'activation et la répression de l'épissage alternatif (Caputi et Zahler, 2002, Min et al., 1995, Chen et al., 1999). On regroupe hnRNP A2/B1 dans la même sous-famille que A1 à raison de leur forte homologie de séquence en acides aminés (68%). Le gène hnRNP A1 code aussi pour l'isoforme A1B, produite par épissage alternatif, résultant en une insertion de 52 a.a. dans le domaine riche en glycines.

HnRNP A1

HnRNP A1 est la plus étudiée des protéines hnRNPs. Elle contient 319 a.a. et a une masse moléculaire de 34 kDa. Étant donné que l'on y retrouve un niveau de transcription important dans un noyau actif, A1 excéderait la concentration nucléaire *in vivo* de 10 μ M ou 7×10^7 copies par cellule en croissance (Michael et al., 1995). La séquence d'acides aminés de A1 est très conservée chez les mammifères, soit 100% de conservation entre l'humain, le rat et la souris, alors que les protéines A1 de l'humain et du *Xenopus laevis* ne diffèrent que de 8% (Kay et al., 1990). HnRNP A1 contient 3 domaines associés à plusieurs activités cellulaires. Deux domaines RRM lui procurent la capacité de lier les acides nucléiques simple brin, ARN ou ADN, et un domaine riche en glycines lui permet d'interagir avec elle-même et les membres de sa sous-famille (Cartegni et al., 1996). Parmi ses activités, on retrouve sa capacité

à moduler la sélection des sites d'épissage 5' (Mayeda et Krainer, 1992) et 3' (Bai et al., 1999), le transport de l'ARNm du noyau au cytoplasme (Izaurralde et al., 1997), la biogénèse des télomères (LaBranche et al., 1998), la stabilité de l'ARN (Hamilton et al., 1997) et le contrôle de la transcription (Hay et al., 2001). La concentration de A1 varie selon le type cellulaire et cette expression différentielle pourrait affecter l'épissage alternatif de plusieurs ARNs pré-messagers (Kamma et al., 1995). Les sections suivantes donnent un portrait de hnRNP A1 selon ses domaines et les fonctions associées à ceux-ci.

RRMs

Les deux premiers domaines de hnRNP A1 sont les domaines RRM1 et RRM2 (**Figure 4A**). Ces motifs ont d'abord été identifiés en comparant les séquences primaires de hnRNP A1 à la poly(A)-binding protein (PABP) (Adam et al., 1986). Un RRM consiste en 80-90 a.a. conservés contenant deux courtes séquences hautement conservées de 6 et 8 résidus, appelé respectivement RNP2 et RNP1 (Dreyfuss et al., 1988). Les deux RRMs sont reliés par une séquence d'environ 15 a.a. soupçonnée d'avoir un rôle capital au niveau de la structure de hnRNP A1. Trois résidus dans ce segment font un contact direct avec l'ADN télomérique; Arg-92, Ser-95, His-101 (Ding et al., 1999). Cette région contribuerait grandement à l'arrangement spatial des deux RRMs lors de la liaison à des séquences d'acides nucléiques. Une étude cristallographique du RRM1 et RRM2 a montré que les deux RRMs pouvaient lier les acides nucléiques par une interaction hélice-hélice antiparallèle entre l'hélice du RRM1 et l'hélice correspondante du RRM2 (Shamoo et al., 1997). Lorsqu'UP1,

protéine correspondant à A1 dont le domaine riche en glycines est absent, lie l'ADN (**Figure 4B**), un changement conformationnel significatif a été observé. Une seconde étude cristallographique est venue appuyer l'observation précédente par la co-cristallisation de deux molécules d'UP1 à deux oligonucléotides d'ADN télomérique simple brin (Ding et al., 1999). Les deux oligonucléotides sont orientés antiparallèlement et chaque extrémité 5' est associée à l'une des protéines via le domaine RRM1, alors que chaque extrémité 3' interagit avec le RRM2 qui se retrouve en face du RRM1 de l'autre protéine du complexe (**Figure 4C**). Le RRM2 agirait comme un stabilisateur du complexe formé en liant le brin opposé. L'étude a également montré qu'UP1 ne fait pas la discrimination entre l'ADN ou l'ARN lors de sa liaison aux acides nucléiques car UP1 ne peut faire aucun contact avec le groupe hydroxyl en position 2' du ribose. Une interaction protéine-protéine contribuerait à la formation du dimère UP1-UP1 en présence d'ADN, à l'aide de 6 a.a. : quatre dans le RRM2 (Ile-164, Lys-166, Tyr-167, His-173) et deux à l'extérieur (Glu-11, Asp-94). On retrouve les interactions suivantes : un lien hydrogène entre Glu-11 et His-173, une interaction électrostatique entre Asp-94 et Lys-166, un lien hydrogène entre Tyr-167 et le groupe carbonyle de Ile-169. De plus, en raison de la symétrie du complexe, chaque interaction est retrouvée deux fois. Néanmoins, un artéfact de compaction des cristaux pourrait expliquer ces observations. Une étude subséquente propose que la conformation de ce complexe UP1 à de l'ADN télomérique serait différente (Vitali et al, 2002).

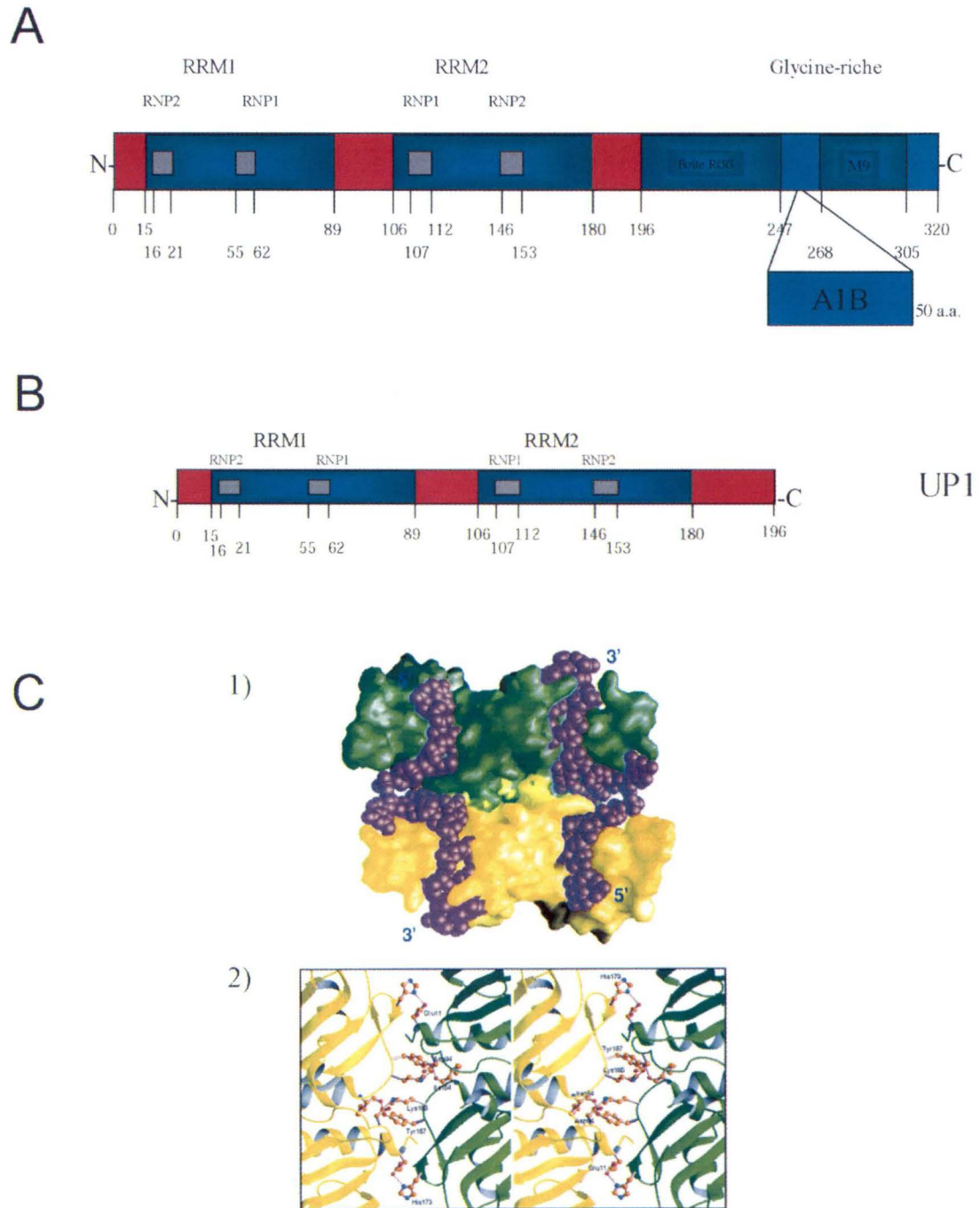


Figure 4. **Structure de hnRNP A1.** A. Les domaines RRM1 et RRM2 sont positionnés respectivement entre les a.a. 15-89 et 106-180. Le domaine glycine-riche s'étend de l'a.a. 196 à 320. Le domaine riche en glycines se compose de la boîte RGG entre les positions 196 et 247, et de la séquence M9 entre les positions 268 et 305. B. Schéma de UP1. C. 1) Arrangement spatial de chaque RRM d'UP1 liant simultanément 2 oligonucléotides positionnés antiparallèlement. 2) Stéréogramme montrant les acides aminés impliqués dans l'interface protéine-protéine. Lien hydrogène (pointillé rouge), atome d'oxygène (rouge), atome d'azote (bleu), atome de carbone (brun) (tiré de Ding et al).

L'ARN et l'ADN simple brin constituent les substrats de hnRNP A1. La séquence spécifique de liaison de A1 aux acides nucléiques a été déterminée par SELEX et correspond à UAGGGA/U avec une constante de dissociation de 3 nM (Burd et Dreyfuss, 1994). Par contre, la séquence optimale équivaut à une duplication de cette séquence espacée de deux nucléotides pour une constante de dissociation de 1 nM. La séquence optimale de liaison pour l'ADN n'a pas été déterminée, mais les répétitions télomériques de vertébrés (TAGGGT) sont très similaires à la séquence d'ARN de liaison optimale pour A1 (LaBranche et al, 1998). Des gels de retardement ont montré qu'A1 n'a pas de difficulté à lier l'ADN simple brin au même titre que la même séquence en ARN (Labrecque, résultats non-montrés). L'étude cristallographique d'UP1 avec 2 oligonucléotides contenant des répétitions télomériques suggère qu'il n'y a pas de contact possible avec le 2'OH du ribose (Ding et al, 1998). L'implication de chaque RRM au niveau de la liaison aux acides nucléiques a été évaluée en présence et en absence du domaine riche en glycines. En son absence, le domaine RRM1 est suffisant pour lier un site de haute affinité, alors que le RRM2 montre une activité de liaison beaucoup plus faible (Dallaire et al., 2000). Par contre, en présence du domaine riche en glycines juxtaposé à un seul RRM, l'activité de liaison du RRM1 serait inhibée, alors que le RRM2 montre une activité de liaison légèrement supérieure (Mayeda et al., 1998). A1 a donc besoin des deux RRMs pour obtenir une activité de liaison de haute affinité et spécifique aux acides nucléiques. Aussi, le RRM2 serait structurellement mieux disposé à la présence du domaine riche en glycines que le RRM1.

Le domaine riche en glycines

Le domaine riche en glycines correspond aux 124 acides aminés de la partie C-terminale de A1 (positions 197 à 320). Il contient 40% de résidus glycine, ce qui rend cette région désorganisée sur le plan structural. Le premier segment de ce domaine (positions 196 à 247) contient quelques répétitions d'a.a. RGG, et est donc très riche en résidus glycine et arginine, ce qui rend cette région très basique. La boîte RGG a été proposée comme un motif de liaison aux séquences d'acides nucléiques simple brin (Kiledjian et Dreyfuss, 1992). Les arginines de cette région seraient sujets à la diméthylation et ceci pourrait modifier l'interaction de A1 aux acides nucléiques (Kim et al., 1997). Une étude sur la protéine hnRNP A2 suggère que cette région est peut-être nécessaire à la localisation cellulaire de la protéine (Nichols et al., 2000). Il existe plusieurs types d'interactions protéine-protéine, en passant de l'interaction ionique au « leucine-zipper ». Dans le cas du domaine riche en glycines de A1, on retrouve plusieurs tyrosines et phénylalanines parmi les résidus glycines et les résidus polaires ou chargés positivement. Ce type de configuration est propice aux interactions hydrophobes. Un essai « pull-down » d'un fragment du domaine riche en glycines comportant des substitutions de 4 a.a. aromatiques consécutifs supporte l'existence de cette interaction (Cartegni et al., 1996).

Malgré les deux RRM, A1 requiert le domaine riche en glycines pour sa liaison coopérative à l'ARN (Cobianchi et al., 1988). Un oligopeptide de 42 a.a., comportant la région boîte RGG, a été synthétisé et contribue à cette propriété *in vitro*. D'autre part, la portion N-terminale (UP1) de A1 contenant les deux RRM, sans le domaine

riche en glycines, est incapable d'entretenir une liaison coopérative à l'ARN (Cobianchi et al., 1988). Sans le domaine riche en glycines, A1 n'est plus retenue sur une colonne d'affinité contenant le domaine riche en glycines couplé à un domaine GST (Biamonti et Riva, 1994). Elle perd aussi son activité à sélectionner l'utilisation des sites d'épissage alternatif 5' distaux en présence de sites de haute affinité (Blanchette et Chabot, 1999). Par rapport à A1, l'isoforme A1B a un domaine riche en glycines plus long. Malgré qu'A1B se lie avec plus d'affinité à l'ARN que A1, A1B a davantage de difficulté à promouvoir la sélection du site d'épissage 5' distal (Mayeda et al., 1994). L'interaction A1-A1 serait ARN-indépendant en s'appuyant sur une expérience d'interaction avec des protéines immobilisées, où l'extrait est préalablement traité à la nucléase micrococcale (Cartegni et al, 1996).

Une séquence de 38 a.a., de la position 268 à 305, a été identifiée comme le signal de localisation nucléaire nécessaire aux imports et exports d'ARNm (Izaurralde et al., 1997). Baptisée M9, cette séquence permet à A1 de faire la navette entre le noyau et le cytoplasme.

HnRNP A1 et l'épissage alternatif

A1 peut influencer la modulation de l'épissage alternatif par différents mécanismes. A1 a initialement été caractérisée comme un répresseur de l'épissage puisqu'elle antagonise l'activité des protéines SR en favorisant l'exclusion des exons alternatifs

et l'utilisation des sites d'épissage 5' distaux. Chez VIH, A1 est capable d'inhiber l'épissage en liant une séquence exonique (ESS) (Caputi et al., 1999) ou intronique (ISS) (Tange et al., 2001). Le mécanisme global consisterait à une répression de la reconnaissance des signaux d'épissage et du site de branchement via la liaison coopérative de A1 au ISS et ESS (Damgaard et al., 2002; Marchand et al., 2002). Deux modèles sont actuellement proposés pour expliquer le mécanisme de répression de A1 au niveau de l'épissage alternatif. Ils seront décrits dans la prochaine section.

La protéine hnRNP A1 est impliquée dans la modulation de plusieurs ARNs pré-messagers. Au cours du développement érythropoïétique, une chute naturelle de la concentration des protéines hnRNP A/B occasionne l'inclusion de l'exon 16 de l'ARN pré-messager codant pour la protéine 4.1R. Ceci s'explique par l'indisponibilité de hnRNP A/B à lier l'ESS situé dans l'exon 16 (Hou et al., 2002). HnRNP A1 régule également son propre ARN pré-messager (Chabot et al., 1997). Le gène contient 11 exons dont le huitième est épissé alternativement afin de produire l'isoforme A1B. La comparaison des séquences introniques autour de l'exon alternatif entre l'humain et la souris a permis d'identifier 10 séquences conservées. Jusqu'à ce jour, les éléments conservés CE1, CE4, CE6 et CE9 ont été caractérisés. CE1 et CE4 ont la capacité de lier les protéines A1 et A2 (Hutchison et al., 2002) **(Figure 5)**. Cette liaison favorise l'exclusion de l'exon 7B sans affecter la reconnaissance du site 5' de cet exon par le snRNP U1. Nous avons établi un modèle pour expliquer l'activité de A1 au niveau de la sélection des sites d'épissage 5'. Nous avons nommé ce modèle « looping-out ». Premièrement, hnRNP A1 lie CE1 et CE4

contenant chacun deux sites de haute affinité pour A1 situés dans la partie intronique flanquant de part et d'autre un exon alternatif. Par une interaction A1-A1, l'exon alternatif se retrouve dans une boucle. Les exons bordant l'exon alternatif sont alors rapprochés, favorisant ainsi l'exclusion de l'exon alternatif (Blanchette et Chabot, 1999). De plus, le site d'épissage 5' se trouvant dans la boucle est réprimé (Nasim et al., 2002). Une récente étude sur le contrôle de l'épissage alternatif de VIH-1 favorise également le modèle du « looping-out ». Dans ce cas-ci, la liaison coopérative de A1 à un ISS et un ESS inhiberait l'épissage en empêchant la reconnaissance de signaux d'épissage essentiels (Damgaard et al., 2002). Un second modèle propose que A1 doit d'abord lier un site de haute affinité et ensuite un recouvrement de l'ARN pré-messager s'amorce par l'assemblage de protéines hnRNP A1 additionnelles, grâce à sa liaison coopérative à l'ARN (Zhu et al., 2001). Ce recouvrement empêche la reconnaissance des signaux d'épissage par le spliceosome. Cependant, ce modèle ne s'applique pas à notre système puisque le snRNP U1 reconnaît toujours le site d'épissage 5' de l'exon 7B.

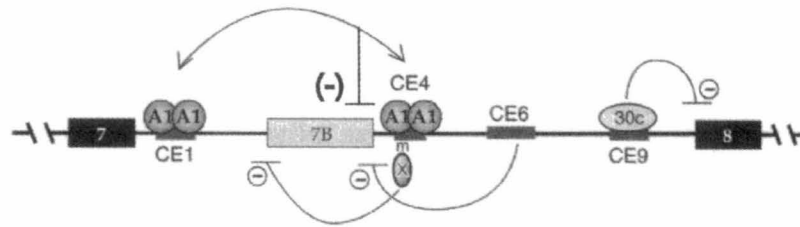


Figure 5. **Mécanismes divers contrôlant l'épissage de l'exon alternatif 7B de l'ARN pré-messager de hnRNP A1.** L'interaction intermoléculaire entre les protéines hnRNP A1 liées à CE1 et CE4 promeut l'exclusion de l'exon alternatif 7B en réprimant l'utilisation du site 5' et en rapprochant les exons 7 et 8. CE6 forme une structure secondaire avec le site d'épissage 5' de l'exon 7B et bloque son utilisation. CE4m empêche l'utilisation du site 3' de 7B. La liaison de SRp30c à l'élément CE9 réprime l'utilisation du site d'épissage 3' de l'exon 8 (tiré de Chabot et al, 2003).

Le mécanisme par lequel la protéine A1 procède, reste néanmoins spéculatif. Sous cette optique, nous avons élaboré ce modèle en tenant compte des caractéristiques de hnRNP A1 connues jusqu'à ce jour. Deux molécules A1 lient d'abord les séquences de haute affinité situées de part et d'autre de l'exon alternatif par l'intermédiaire du domaine RRM1, puisqu'il possède à lui seul une activité de liaison forte et spécifique aux acides nucléiques (Dallaire et al., 2000) (Figure 6). Quoiqu'une expérience indique que l'interaction A1-A1 est indépendante de l'ARN (Cartegni et al., 1996), certaines indications nous suggèrent que la liaison d'A1 à l'ARN stimule l'interaction A1-A1 (G. Lettre et M. Blanchette, résultats non-publiés). Il y aurait donc interaction protéine-protéine à l'aide du domaine riche en glycines entre les protéines A1 situées de chaque côté de l'exon alternatif. La liaison du RRM1 à l'ARN et l'interaction protéine-protéine du domaine riche en glycines occasionneraient un changement de conformation découvrant le RRM2. Enfin, le RRM2, dont le rôle dans l'interaction protéine-protéine n'est pas exclu, servirait à stabiliser le complexe en liant la séquence ARN orientée antiparallèlement, amenée par l'interaction A1-A1 (Figure 6).

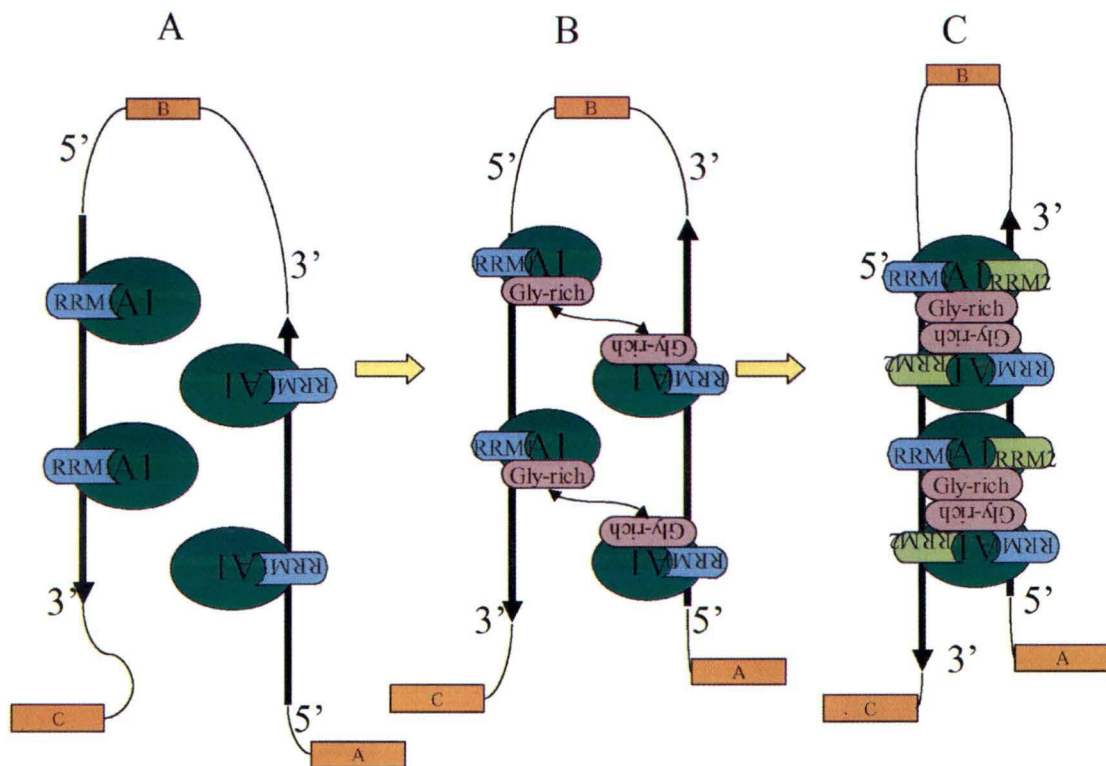


Figure 6. **Modèle : activité de hnRNP A1 dans l'épissage alternatif.** A. Liaison du RRM1 à une séquence d'ARN de haute affinité. B. Rapprochement des deux séquences d'ARN à l'aide d'une interaction protéine-protéine par le domaine glycine-riche de A1. C. Stabilisation du complexe par la liaison du RRM2 à l'ARN.

Dans le but de vérifier ce modèle, nous avons approché génétiquement la problématique en tentant d'identifier *in vivo* des mutants ayant une déficience au niveau de la liaison coopérative de molécules de A1. Un criblage génétique basé sur la répression traductionnelle chez *E.coli* a été utilisé pour identifier des régions dans les domaines RRM2 et riche en glycines impliquées dans l'interaction A1-A1. Nous avons confirmé que la plupart des mutants sont affectés au niveau de l'interaction protéine-protéine par la technique de double-hybride chez la levure. Parmi les mutants A1 récupérés, certains ont une activité réduite au niveau de la liaison à l'ARN qui interfère avec l'activité de A1 dans l'épissage. Le séquençage de nos mutants révèle que plusieurs régions situées dans le domaine riche en glycines contribuent à l'interaction A1-A1. Des mutations au niveau des acides aminés dans le RRM2 affectent l'interaction A1-A1, mais perturbent également la liaison de A1 à l'ARN. D'autres mutations affectent peut-être la structure, la stabilité et le niveau d'expression de la protéine. Nous vérifions la stabilité et le niveau d'expression de la protéine par immunobuvardage. Nous ne pouvons écarter, hors de tout doute, les mutants ayant subi une modification de la structure de la protéine. Cependant, nous jugeons que la structure doit être modifiée que très localement lorsque les protéines mutantes sont aptes à lier les acides nucléiques au même niveau que la protéine wt. Par gel de retardement, il est impossible de départager entre la possibilité que le défaut dans la liaison à l'ARN entraîne une réduction de l'interaction A1-A1, ou si cette mutation est importante pour l'interaction A1-A1 indépendamment de l'effet sur la liaison à l'ARN. Enfin, puisque la liaison coopérative de A1 à l'ARN nécessite une interaction protéine-protéine, cela devrait se refléter dans l'activité de A1 dans

l'épissage alternatif. Selon le modèle « looping-out », l'interaction A1-A1 est nécessaire pour rapprocher les sites d'épissage distaux. Nous avons alors testé l'activité de nos mutants A1 dans la sélection de sites d'épissage 5' *in vitro* dans un extrait nucléaire déplété en hnRNP A1. La force d'interaction A1-A1 semble influencer A1 dans sa capacité à promouvoir l'utilisation des sites d'épissage 5'. Donc, les acides aminés participant à l'interaction A1-A1, influencent la protéine hnRNP A1 dans son activité sur l'épissage alternatif.

MATÉRIEL ET MÉTHODES

Constructions plasmidiques

pGEX-A1 : Le plasmide pGEX-A1 a été construit en insérant l'ADNc de A1 (Ben-David et al., 1992) au site Eco RI de pGEX-2T. Le fragment UP1 (LaBranche et al., 1998) a été introduit au même site de pGEX-2T. pGEX-A1 Δ RRM2 provient de la délétion du fragment Bgl II-Msc I et contient une séquence additionnelle contenant un site de restriction unique. A1 Δ RRM2 a été construit à partir du vecteur pGEX-KG aux sites Bam HI et Eco RI.

pRA1 : pRA1 correspond au vecteur pREV1 où une librairie d'ADNc hnRNP A1, contenant des mutations dans le domaine RRM2 et riche en glycines, a été insérée aux sites Bsu36 I et Nde I remplie préalablement à la Klenow.

pUP2A1BS et pUP1A1BS : les oligonucléotides
5'AAAGGTACCTATGATAGGGACTTAGGGTGNNNNNNNNNNNNNTATG
ATAGGGACTTAGGGTGAGACAACAAGATGTGCGAAC 3', et
5'AAAGGTACCTAGGGTNNNNNNNNNNNNNNNNNNNTAGGGTAGACAA
CAAGATGTGCGAACTC 3' sont clonés dans le vecteur pUPhpII aux sites Kpn I et Hind III.

pQE A1 : L'ADNc de A1 de type sauvage et de la librairie de mutants ont été insérés dans le plasmide pQE-80L (Qiagen) aux sites Bgl II et Hind III. Les fragments proviennent du plasmide pRA1 digéré aux mêmes sites de restriction. Tous les clones ont été séquencés.

pGAD-A1 : pGAD A1 et pGAD UP1 ont été construits par Catherine LeBel. A1- Δ RRM2 a été inséré dans pGAD GH (Clontech) aux sites Bam HI et Eco RI à partir du plasmide pGEX A1- Δ RRM2. Les mutants A1 sélectionnés suite à l'expérience de répression traductionnelle ont été insérés dans pBluescript II KS(+) aux sites Bam HI et Hind III, à partir du fragment isolé des plasmides pQE-A1 mutants. Ensuite, les mutants A1 ont été digérés et transférés dans le vecteur pGAD GH aux sites Bam HI et Xho I.

pAS2-A1 : pAS2 A1 et pAS2 UP1 ont été construits par Catherine LeBel. D'abord, le vecteur pAS2-1 (Clontech) a été modifié afin d'arriver dans le bon cadre de lecture en remplissant le site Nde I à l'aide de la Klenow. À partir du vecteur pBluescript KS(+) A1, les mutants A1 ont été insérés aux sites Bam HI et Sal I du vecteur pAS2-1.

Mutagenèse par PCR

La mutagenèse aléatoire a été effectuée sur la construction plasmidique pRA1 de manière à ce que chaque molécule contienne environ une mutation. Le travail a été effectué par Martin Simard.

Répression traductionnelle

Le travail a été effectué par Martin Simard, Daniel Garneau et Timothée Revil. La librairie pRA1 est transformée dans les souches *E.coli* WM1/F' (Celandier, 2000) contenant le vecteur pLACZ. Les colonies sont induites avec 600 μ M d'IPTG pendant 4 heures à 37°C afin d'exprimer les protéines hnRNP A1. Les colonies blanches (perte d'interaction A1-A1) sont sélectionnées. Les cellules sont lysées par sonication et l'expression des protéines est vérifiée par Western blot à l'aide de l'anticorps anti-A1 #15 (Hutchison et al., 2002). La liaison des mutants de hnRNP A1 aux acides nucléiques est vérifiée par la technique de gel de retardement en incubant l'extrait bactérien total avec l'oligonucléotide Tel-3 (Blanchette et Chabot, 1999). Tel-3 contient une séquence d'acide nucléique constituée de 3 répétitions d'ADN télomériques humaines (TTAGGG).

Transcrits radiomarqués pour réaction d'épissage *in vitro* et gel de retardement

Les constructions plasmidiques utilisées pour la production de transcrits radiomarqués ont été élaborées par un collègue du laboratoire : C5' 4/4 (Blanchette et Chabot, 1999). L'ADN plasmidique a été récupéré soit par la méthode de centrifugation au chlorure de césium, soit par lyse alcaline (Sambrook, 1989). Les plasmides ont été linéarisés à l'aide de l'enzyme de restriction Sca I (NEB) pour C5' 4/4. L'ARN polymérase T3 a été utilisée pour la transcription. Les ARNs sont produits en présence d'un analogue de CAP et incorporés d'UTP radiomarqué (UTP α -³²P) (Amersham Pharmacia, NEN).

Marquage d'oligonucléotides pour gel de retardement

TS10 (Blanchette et Chabot, 1999) constitue la séquence d'oligonucléotides marqués. Cette séquence contient 9 répétitions d'ADN télomériques humaines (TTAGGG). Les oligonucléotides sont phosphorylés à l'extrémité 5' à l'aide du [$\gamma^{32}\text{P}$]ATP (Amersham) sous l'action de la T4 polynucléotide kinase (NEB) (Sambrook, 1989). Les produits sont purifiés sur une colonne MicroSpin G-25 (Pharmacia).

Déplétion d'extraits nucléaires en hnRNP A1

La préparation des extraits nucléaires HeLa est décrite dans (Dignam et al., 1983). Les extraits sont incubés à deux reprises, 10 min à 37°C en présence d'un tampon d'épissage et de billes couplées à 50 nM de l'oligonucléotide Tel-3 (Blanchette et Chabot, 1999). L'efficacité de la déplétion de l'extrait nucléaire est vérifiée par immunobuvardage avec l'anticorps anti-A1 clone #13.

Détection de hnRNP A1 par immunobuvardage (Western blot)

Les protéines sont séparées par électrophorèse sur un gel de polyacrylamide SDS-PAGE 12.5%, et ensuite transférées sur une membrane Hybond-C (Amersham). La membrane est bloquée non-spécifiquement dans un tampon TBS-Tween pH 7.6 (10 mM Tris, 150 mM NaCl 0.1% Tween) additionné de lait en poudre (Carnation) dans

une proportion de 10%. Suite à 3 lavages de 10 minutes du tampon TBS-Tween, la membrane est incubée 1 h à 4°C avec l'anticorps anti-A1 clone #13 (dilué 1 dans 1000 dans le TBS-Tween) reconnaissant hnRNP A1/A1B, A2 et B1. Toujours suite à 3 lavages de 10 minutes du tampon TBS-Tween, la membrane est réincubée 1 h à température ambiante avec l'anticorps secondaire anti-lapin (Amersham Pharmacia Biotech) reconnaissant la partie constante de l'anticorps de lapin. La présence des protéines est révélée à l'aide des réactifs chemiluminescents ECL (Amersham Pharmacia Biotech). La membrane est exposée 30 sec ou 210 sec sur un autoradiogramme XAR (Kodak).

Production et purification de protéines recombinantes

Toutes les protéines recombinantes hnRNP A1 et ses mutants sont exprimées dans *Escherichia coli* BL21. Les bactéries croissent 2-3 heures jusqu'à ce qu'elles atteignent une densité entre 0.5 et 1.0 à 600 nm sur le spectrophotomètre. Une induction avec 100 µM IPTG est suivie d'une incubation de 4 heures à 37°C pour A1 étiquetée GST, alors que la protéine avec une étiquette 6xHis est induite avec 50 µM IPTG, 50 µM PMSF et incubée pendant 1 heure. Les cellules sont lavées avec du Rec-buffer (50 mM pipérazine-HCl pH 9.8, 500 mM NaCl, 1 mM DTT, 1 mM EDTA, 0.5 mM PMFS, 1 mM benzamidine, 20 µg/µl bacitracine) et resuspendues avec 10 ml de Rec-Buffer par 100 ml de culture. Le lysosyme (0.4 mg/ml) est ajouté 30 min avant la sonication. Les cellules sont soniquées sur glace pendant 10 sec, à 3

reprises. Le Triton X-100 1% est ajouté et le tout est incubé 30 min à 4°C sur une plaque rotative (Nutator). Par la suite, l'extrait cellulaire est centrifugé 10 min à 10 000 rpm à 4°C et le surnageant est déposé sur une colonne de 600 µl de glutathione-Sephadex (50% dans Rec-buffer). Le mélange est incubé 30 min à 4°C, centrifugé et lavé 4 fois avec du Rec-buffer 0,1% Triton X-100. Les protéines-GST sont éluées à l'aide de 300 µl de glutathione-réduit, 200 mM pipérazine pH 9.8, 1 mM DTT, 1 mM EDTA, et 500 mM NaCl. Suite à l'induction, les protéines étiquetées 6xHis sont resuspendues dans un tampon de lyse pH 8 (300 mM NaCl, 50 mM NaH₂PO₄, 50 mM pipérazine-HCl pH 9.8, 0.5 mM PMFS, 1 mM benzamidine, 20 µg/µl bacitracine, 1 mM imidazole). On ajoute le lysozyme (1 mg/ml) et incube 30 min sur glace avant de soniquer 10 sec, à 3 reprises. Les protéines étiquetées 6xHis sont purifiées sur des colonnes Ni-NTA (Qiagen), selon le protocole décrit (Qiagen, 2001). Le tampon d'éluion contient également 50 mM pipérazine-HCl pH 9.8, 0.5 mM PMFS, 1 mM benzamidine, 20 µg/µl bacitracine. Toutes les protéines recombinantes éluées sont dialysées dans du tampon D (20 mM Hepes-KOH pH 7.9, 100 mM KCl, 20 % glycérol, 0.2 mM EDTA, 1 mM DTT, 0.5 mM PMSF). La concentration des protéines recombinantes a été mesurée par Bradford (Bio-Rad) en comparant une série de dilutions de BSA (bovine serum albumin) comme standard, et vérifié par immunobuvardage avec la concentration connue d'une protéine.

Gel de retardement

Les essais de liaison des protéines recombinantes His-A1 et GST-A1 et leurs mutants ont été effectués avec l'oligonucléotide TS10, dans un tampon d'épissage incubé 30 min à la température de la pièce. Les réactions sont réincubées 5 min en présence d'héparine (0.4 µg/µl) comme compétiteur non-spécifique. Les réactions sont fragmentées sur un gel 5% polyacrylamide (29 :1 acrylamide / bisacrylamide) non-dénaturant dans un tampon Tris-glycine (50 mM Tris pH 8.8, 50 mM glycine). Les signaux ont été quantifiés directement par autoradiographe électronique (Packard Bioscience Company) et quantifiés en utilisant le logiciel « Quantity One » après exposition sur film XAR (Kodak).

Réaction d'épissage *in vitro*

Les transcrits ont été incubés 2 h à 30°C dans un extrait nucléaire de cellules HeLa en présence d'un tampon d'épissage (Chabot et al., 1997) ou dans un extrait de cellules HeLa déplété en hnRNP A1. Les ARNs épissés ont migré sur un gel de polyacrylamide (38 :2 acrylamide/bisacrylamide) 11% contenant 8 M urée, et sont quantifiés par autoradiographe électronique (Packard Bioscience Company).

Technique double-hybride

Technique effectuée par Catherine LeBel. Les plasmides pGAD-A1 et ses mutants ainsi que pAS2-A1 et ses mutants, ont été co-transformés dans la souche de levure PJ69-4a (James et al., 1996) et ont été sélectionnés sur un milieu pauvre en leucine et tryptophane à 30°C. La présence d'interaction entre les protéines A1 est évaluée par la capacité des levures à croître sur des milieux déficients en adénine, et en histidine, en plus d'un test d'expression de la β -galactosidase. La force d'interaction entre les protéines A1 est mesurée à l'aide de l'expression de la β -galactosidase en milieu liquide (en unité Miller) considérant la quantité de cellules et le temps de la réaction (Ausubel, 1998). Ce travail a été effectué conjointement avec Aline Simoneau.

RÉSULTATS

Selon le modèle du « looping-out », nous croyons que le domaine riche en glycines serait nécessaire à l'interaction entre protéines hnRNP A1. Une analyse précédente indique la contribution essentielle du domaine riche en glycines. En effet, un mutant de A1 appelé UP1, manquant les 124 derniers a.a. en queue C-terminale, peut lier l'ARN mais a perdu l'activité de modulation de l'épissage alternatif de l'ARN pré-messager de hnRNP A1 (Blanchette et Chabot, 1999). Basé sur les données cristallographiques du complexe UP1 à l'ADN télomérique, nous suspectons également le domaine RRM2 de participer à l'interaction A1-A1. Nous avons donc adopté une approche génétique afin d'identifier avec plus de précision les régions et les acides aminés contribuant à l'interaction A1-A1.

Il faut d'abord mentionner que ce projet a été réalisé avec la participation de plusieurs personnes. Martin Simard a mis au point le système de répression traductionnelle et effectué la mutagenèse. Daniel Garneau et Timothée Revil l'ont accompagné dans le criblage des mutants, l'élimination des faux-positifs en vérifiant l'expression des protéines chez la bactérie et leur capacité de liaison aux acides nucléiques en analysant l'extrait cellulaire total. Le projet me fut confié par la suite. J'ai procédé à tous les clonages, participé au séquençage, exprimé, purifié et quantifié les protéines recombinantes pour ensuite les caractériser au niveau de leur liaison aux acides nucléiques, mesuré leur activité dans l'épissage alternatif. J'ai également participé à l'évaluation et la quantification de la force d'interaction protéine-protéine par la

technique de double-hybride avec Aline Simoneau et Catherine LeBel qui a mis au point ce système.

Détection de l'interaction A1-A1 par répression traductionnelle

Un essai de répression traductionnelle a été développé dans le but d'étudier la liaison de protéine à son substrat ARN (Jain et Belasco, 1996). La stratégie consiste à placer la séquence d'ARN cible en amont de la séquence Shine-Dalgarno sur un ARNm rapporteur (**Figure 7A**). En absence de la protéine, l'ARN de la sous-unité 30S du ribosome bactérien est recruté par la séquence Shine-Dalgarno et l'ARNm est traduit. Dans le cas où la protéine est exprimée et lie la séquence cible, un empêchement stérique perturbe le recrutement de la sous-unité ribosomale et interfère avec la traduction. Ce système, combiné à la mutagenèse de la protéine sous étude, permet d'identifier les régions importantes contribuant à la liaison à l'ARN.

Le système a été modifié selon un système de répression traductionnelle spécialement établi pour étudier la liaison coopérative de 2 molécules pouvant lier l'ARN (Jain et Belasco, 2001). D'abord, on s'est assuré que la liaison de deux protéines A1, immédiatement en amont de la séquence Shine-Delgarno, ne soit pas suffisante pour provoquer une répression traductionnelle (**Figure 7B**) (Travaux réalisés par M. Simard en collaboration avec C. Jain et J. Belasco, Skirball Institute, N.Y.). Ceci indique que la répression ne serait pas causée par la masse des molécules A1 situées à

proximité de la séquence d'initiation de la traduction. Deux autres sites de haute affinité pour hnRNP A1 ont été insérés en amont des premiers sites, séparés par une séquence de 20 nucléotides. La répression traductionnelle est obtenue lorsque les 4 sites sont présents, suggérant qu'une liaison coopérative serait essentielle pour qu'il y ait répression (**Figure 7B**). Cette configuration serait représentative de la dimérisation de A1, nécessaire à la sélection des sites 5' d'épissage alternatif en présence de sites de haute affinité de part et d'autre d'un exon alternatif. Lorsqu'A1 est remplacée par UP1, on n'obtient aucune répression de la traduction, signifiant que le domaine riche en glycines est nécessaire à la liaison coopérative de A1.

Ainsi, nous évaluerons les a.a. situés dans le domaine riche en glycines et le RRM2 potentiellement impliqués au niveau de l'interaction protéine-protéine par la récupération des protéines mutantes suite à la répression traductionnelle.

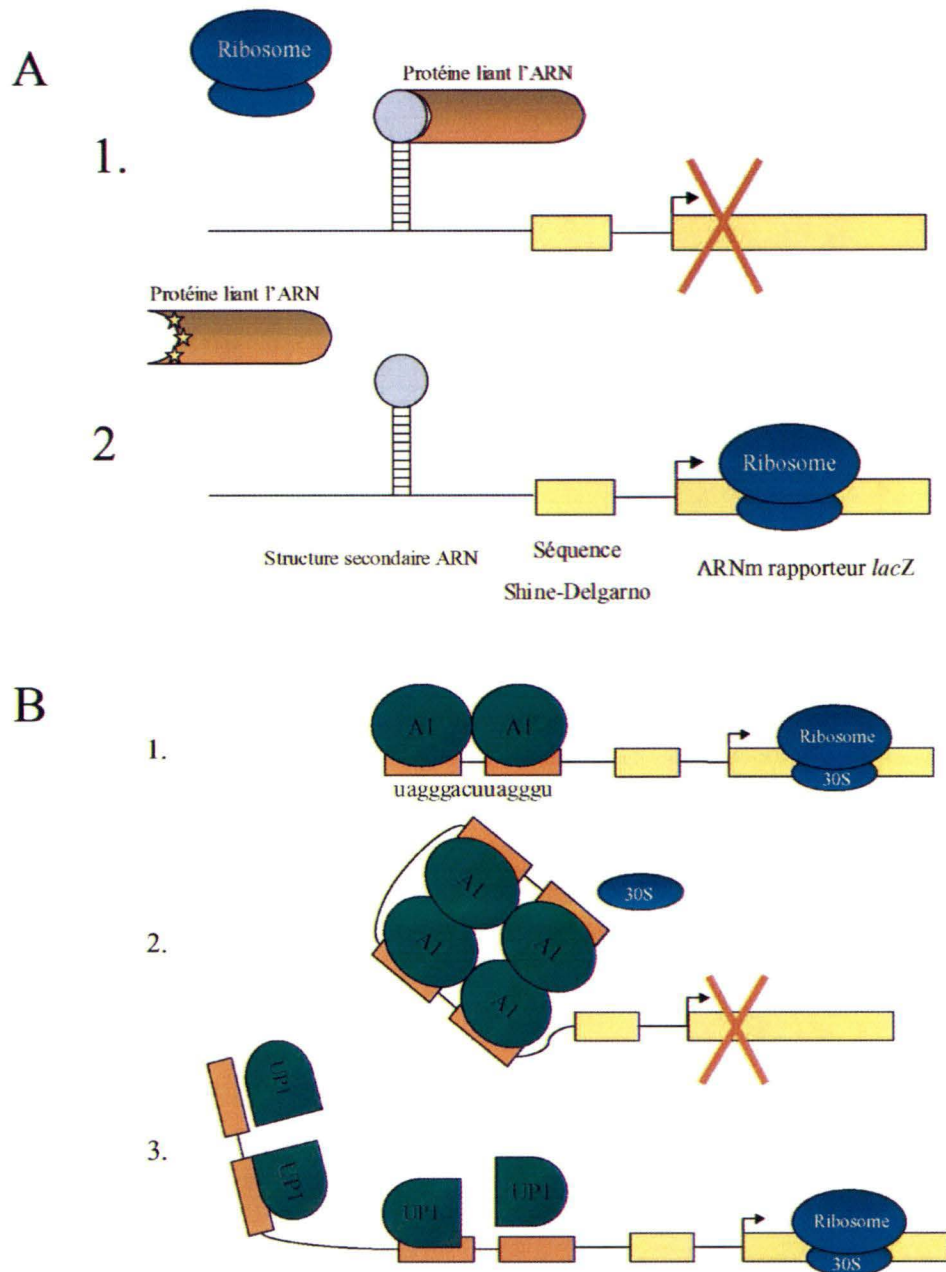


Figure 7. **Répression traductionnelle.** A. L'interaction d'une protéine liant l'ARN à sa séquence ou structure cible en amont de la séquence Shine-Delgarno, empêche stériquement la reconnaissance du ribosome à ce site, donc l'initiation de la traduction (1). La traduction de l'ARNm rapporteur, suite à une mutagenèse de la protéine ou de la séquence cible d'ARN, indiquera que la mutation a affecté la liaison de la protéine à sa séquence ou structure (2). B. Similairement à la répression traductionnelle traditionnelle, la liaison coopérative de hnRNP A1 à une séquence ARN de haute affinité empêche la traduction de l'ARNm rapporteur (2). La présence d'une seule séquence ARN de haute affinité (1) et l'absence du domaine riche en glycines de A1 permet la traduction de l'ARNm rapporteur (3).

Une mutagénèse aléatoire dirigée dans les domaines RRM2 et riche en glycines de l'ADNc de hnRNP A1 par la technique d'amplification par erreur de la polymérase par PCR a été réalisée par M. Simard. La *Taq* polymérase génère environ un changement par 9000 nucléotides sous des conditions normales, pouvant modifier la nature d'un a.a. (Innis, 1990). Par la suite, la librairie nouvellement créée a été transformée dans une souche *E.coli* WM1/F' (Celandier, 2000). Cette souche contient un mutant du gène *recA* permettant d'augmenter la stabilité des plasmides transformés et d'augmenter la reproductibilité des tests de β -galactosidase. La souche contient également le gène du répresseur *lac* contrôlant la synthèse de hnRNP A1 dans le cas où l'expression excessive du mutant A1 serait toxique pour la cellule.

LacZ est l'ARNm rapporteur utilisé pour le test de répression traductionnelle. En principe, la présence d'une interaction A1-A1 empêche la reconnaissance du site d'initiation par le ribosome et on le visualise par l'absence d'activité de la β -galactosidase sur des plaques IPTG/XGAL. Aucune coloration ne sera détectée lorsque la protéine A1 de type sauvage est exprimée. À l'opposé, l'absence d'une interaction A1-A1 provoquera la traduction de *lacZ* et des colonies bleues apparaîtront. Ensuite, les mutants repêchés ont été testés dans une seconde souche WM1/F', et la majorité ont obtenu un phénotype bleu ou bleu pâle. Les protéines ont été extraites de ces colonies et un immunobuvardage de type Western a été effectué afin de vérifier le niveau d'expression et la stabilité de la protéine A1. Les faux-positifs causés par un faible niveau d'expression de A1 ont été éliminés tout comme les protéines instables. Un gel de retardement de l'extrait cellulaire total a également

été effectué afin d'écarter tous les candidats défectueux dans la liaison avec un site de haute affinité pour A1. Finalement, les candidats restant ont été séquencés.

Séquences

Le séquençage a révélé qu'il y avait en moyenne deux à trois mutations par clone, ce qui constitue un nombre de mutations par protéine plus élevé que prévu (**Figure 8**). Certains phénotypes des mutants A1 pourraient provenir de la combinaison de mutations. Puisque la purification des protéines recombinantes a été effectuée en même temps que le séquençage, les clones contenant des mutations multiples n'ont pas été écartés. Ainsi, on peut vérifier si tous les candidats ressortis par le système de répression traductionnelle montrent les mêmes caractéristiques de liaison aux acides nucléiques, d'interaction protéine-protéine et d'activité au niveau de l'épissage alternatif. On dénombre certaines mutations dans le RRM2, mais aucun mutant de A1 contenant un des 4 a.a. mentionnés dans l'étude cristallographique participant à l'interaction UP1-UP1 (Ding et al., 1999) n'a été repêché. Plusieurs changements d'a.a. et même quelques troncatures ont été répertoriés à divers endroits dans le domaine riche en glycines. Cependant, aucune corrélation ne peut réellement être établie entre les mutants de A1 puisqu'il y a souvent plus d'une mutation par protéine.

La liaison aux acides nucléiques est affectée par certaines mutations

Nous avons préalablement écarté les candidats repêchés par la stratégie de répression traductionnelle qui montraient une déficience dans la liaison avec les acides nucléiques. Cette technique utilisait un extrait bactérien brut et ne permettait pas d'évaluer avec précision la force d'interaction de chaque mutant A1 par rapport à la protéine A1 wt. Nous avons donc voulu confirmer la force d'interaction des mutants A1 par gel de retardement à l'aide de l'oligonucléotide TS10 contenant 9 répétitions d'ADN télomérique. En premier lieu, il a fallu produire nos A1 mutants sous forme de protéines recombinantes. Dans les faits, cette étape fut accomplie avant de connaître l'identité des mutations de chaque clone. L'insertion d'une étiquette 6xHis en position N-terminale a été la stratégie choisie à raison de sa petite taille suggérant une faible interférence avec la structure et la fonction de la protéine. De plus, la purification peut se faire grâce à la liaison spécifique de l'étiquette 6xHis à une colonne de nickel. L'état des protéines recombinantes purifiées a été vérifié par immunobuvardage à l'aide de l'anticorps anti-A1 clone #13 (**Figure 9**). Certains mutants A1 sont moins bien exprimés dans les bactéries BL21 (piste #4). Pratiquement toutes les préparations ont une bande additionnelle légèrement plus courte, probablement causée par la dégradation pendant l'expression ou la purification de la protéine, ou bien encore par un artefact de migration provoqué par un repliement différent de la protéine. Quelques mutants ont une migration plus rapide que le His-A1 wt (pistes #5 et 8).

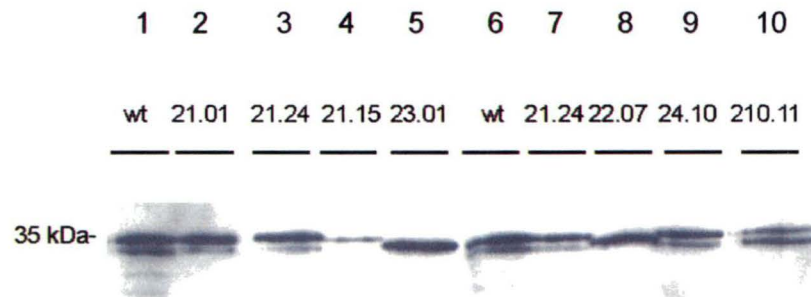


Figure 9. États des protéines recombinantes hnRNP A1 et ses mutants.

Immunobuvardage des protéines His-A1 mutantes en comparaison avec le His-A1 wt. Environ 12 pmoles de chaque protéine recombinante purifiée ont été déposées sur un gel SDS-PAGE 12.5%, sauf pour la piste #4 où 2.5 pmoles ont été déposées. Un contrôle His-A1 de type sauvage (piste #1) permet d'évaluer le poids moléculaire et le niveau de dégradation des protéines recombinantes. Les protéines ont été détectées à l'aide d'un anticorps anti-A1.

Par la suite, les protéines recombinantes ont été incubées 30 min à la température de la pièce en présence de l'oligonucléotide TS10. L'ajout d'héparine, agissant comme compétiteur, réduit les interactions non-spécifiques (Figure 10). Le taux de disparition de l'ADN libre en fonction de la concentration de hnRNP A1 a été calculé (résultats non-montrés), et le profil de migration a été visuellement considéré. Les mutants 210.11, 23.01, 24.10 et 21.14 ont une activité de liaison à l'ADN équivalente au His-A1 de type sauvage. Les mutants 21.15, 21.24 et 22.07 ont une activité réduite. Les candidats ayant une activité équivalente à A1 sont retenus pour évaluer leur activité au niveau de l'épissage alternatif.

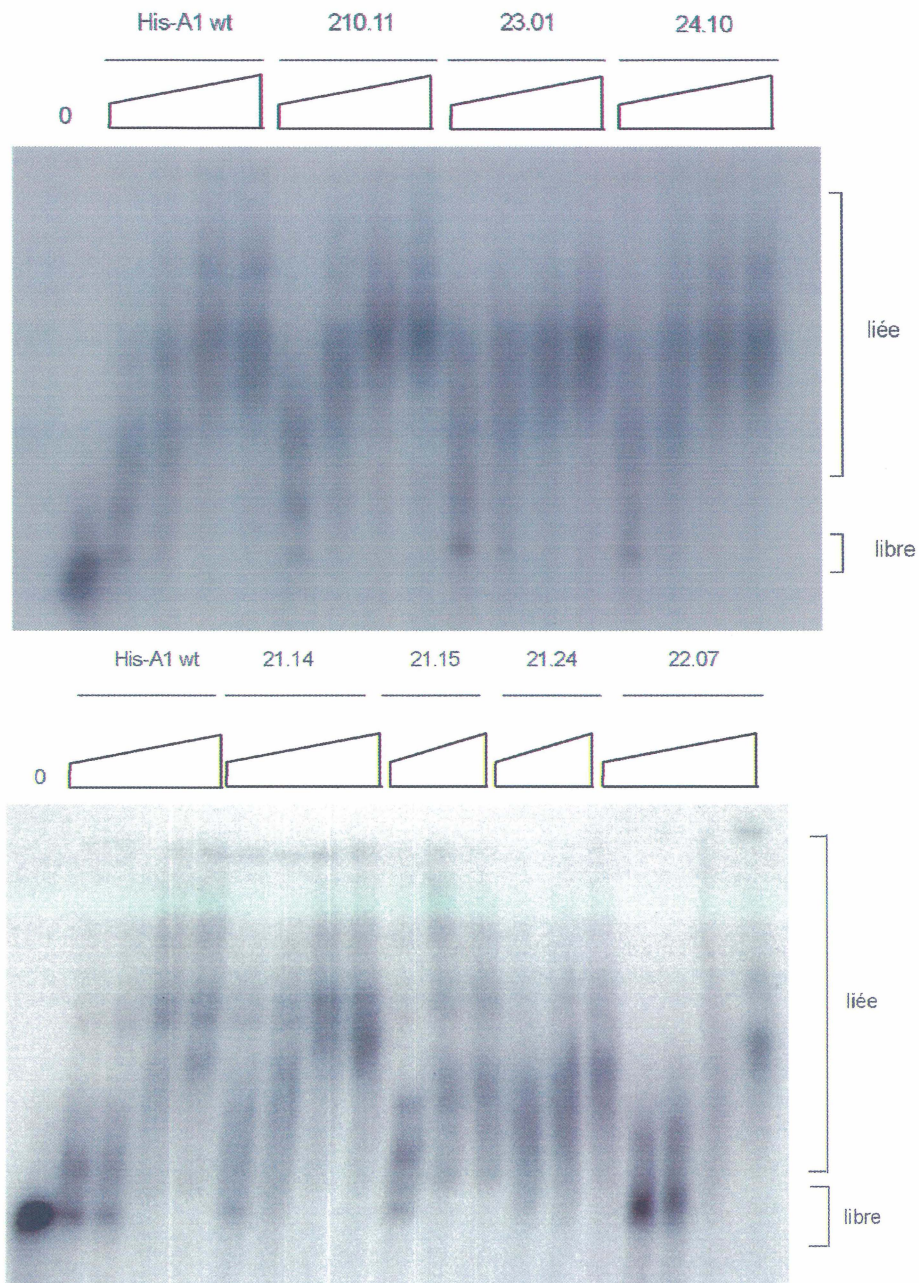


Figure 10. Affinité des protéines hnRNP A1 mutantes à l'ADN télomérique simple brin..

Comparaison de la liaison d'ADN télomérique simple brin à des concentrations croissantes (0, 0.20, 0.40, 0.96 et 1.92 pmol/ul) de la protéine recombinante His-A1 wt versus une concentration croissante de chaque mutant His-A1. Les protéines recombinantes ont été incubées 30 min à la température de la pièce en présence d'héparine, un compétiteur non-spécifique. Les échantillons ont été séparés sur un gel polyacrylamide 5% non-dénaturant.

Interaction A1-A1 des mutants selon le double hybride chez la levure

Les mutants repêchés par répression traductionnelle devraient, par définition, être déficients dans l'interaction A1-A1. L'interaction protéine-protéine des mutants repêchés a été confirmée par une autre approche, soit le double hybride chez la levure (réalisation en collaboration avec Catherine LeBel et Aline Simoneau). Il a déjà été montré en double hybride que l'absence du domaine riche en glycines entraîne l'incapacité d'A1 d'interagir avec elle-même (Cartegni et al., 1996). Nos protéines mutantes ont été fusionnées au domaine de liaison à l'ADN Gal4 et une autre série a été fusionnée au domaine de transactivation Gal4. Ainsi, l'interaction de chaque mutant est testée avec lui-même. La présence d'interaction a été détectée en triplicata par la croissance des levures en milieu restrictif en adénine (**Figure 11A**). Comparés au A1 wt, les mutants UP1, 210.11 et 21.07 montrent une activité très réduite alors que les autres mutants, 21.14, 23.01, et 24.10 semblent peu affectés. Cependant, il est difficile d'évaluer l'interaction entre protéines sur un milieu restrictif, car il est plus difficile de distinguer entre une interaction négative, un problème de croissance, ou une faible quantité de levures appliquées. Afin d'être plus quantitatif, la force d'interaction des protéines est mesurée en triplicata par un test de la β -galactosidase en milieu liquide. Ce test mesure, en unité Miller, le taux de transcription du gène rapporteur *lacZ* en fonction de la quantité de cellules, du temps et du volume de la réaction (**Figure 11B**). L'expression des protéines chez la levure n'a pas été vérifiée, excepté pour les contrôle hnRNP A1 wt et UP1. On a supposé que le niveau

d'expression serait semblable au niveau d'expression chez la bactérie, malgré que le système soit différent et que les modifications post-traductionnelles peuvent affecter différenciellement l'expression. Donc, les résultats négatifs ne sont pas réellement significatifs. Les tests de β -galactosidase en milieu liquide suggèrent que tous les mutants ont une déficience au niveau d'une interaction protéine-protéine par rapport au A1 de type sauvage. Seuls les mutants UP1 et 21.07 ont perdu totalement la capacité d'interagir avec eux-mêmes. Il est possible que l'ARN participe à l'interaction A1-A1 chez la levure, puisque l'interaction A1-A1 serait favorisé lorsqu'A1 lie les acides nucléiques. Néanmoins, UP1 nous montre qu'il ne peut interagir avec lui-même dans un test double-hybride alors que nous savons qu'il est capable de très bien lier les acides nucléiques *in vitro* suite à nos résultats sur gel de retardement.

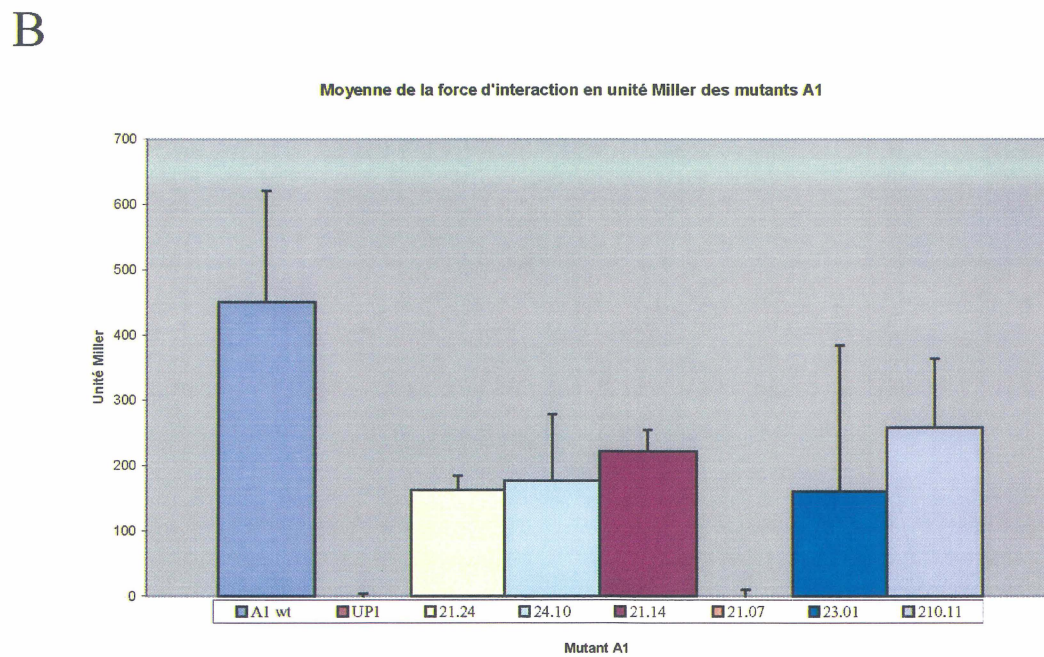
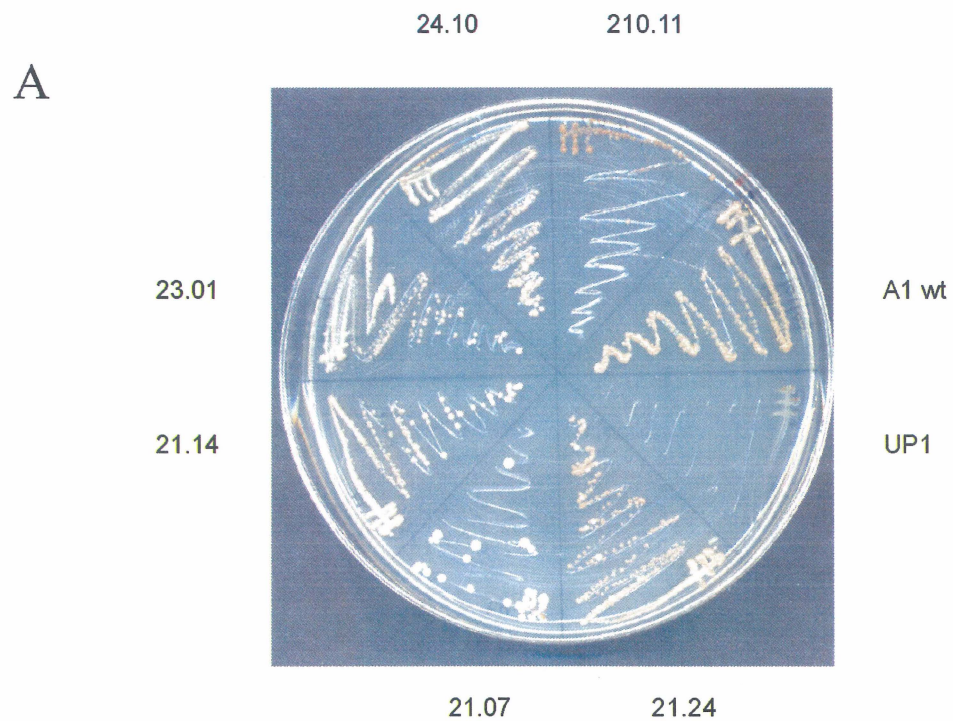


Figure 11. **Mutants hnRNP A1 en double hybride.** A. Croissance des levures sur un milieu restrictif en adénine. B. Tableau de la quantification de la force d'interaction des mutants hnRNP A1 par les réactions Miller.

Certaines mutations dans le domaine riche en glycines affectent la sélection des sites d'épissage 5'

L'absence du domaine riche en glycines empêche A1 de promouvoir la sélection vers le site 5' distal d'un transcrit contenant deux sites d'épissage 5' pour un site 3' (Blanchette et Chabot, 1999). Une interaction protéine-protéine serait donc nécessaire pour maintenir l'activité de hnRNP A1 dans l'épissage alternatif. Les mutants repêchés par la technique de répression traductionnelle, basée sur une perturbation de l'interaction A1-A1, devraient donc avoir une activité réduite au niveau de l'épissage alternatif. Afin de vérifier ceci, nous avons d'abord déplété A1 d'un extrait nucléaire d'épissage et ensuite, supplémenté cet extrait déplété avec les protéines A1 mutées. Un extrait nucléaire de cellules HeLa est déplété des protéines A1/A1B, A2/B1 en l'incubant pendant 20 min à 30 °C sur une colonne de billes d'agarose où des oligonucléotides d'ADN contenant 3 répétitions télomériques simple brin y sont couplés (**Figure 12A**). L'efficacité de la déplétion de hnRNP A1 endogène est vérifiée par immunoblot à l'aide d'un anticorps anti-A1 (**Figure 12B**). L'efficacité de déplétion est de l'ordre de 85-95% (pistes #3 et #4). Nous avons utilisé l'ARN pré-messager modèle contenant deux sites d'épissage 5' en compétition pour un site 3', et quatre sites de haute affinité pour A1 de part et d'autre d'un exon alternatif (Blanchette et Chabot, 1999) (**Figure 12C**). Le transcrit marqué au ^{32}P a été incubé dans l'extrait nucléaire préalablement déplété en A1 dans lequel les protéines recombinantes A1 sont ajoutées en quantité croissante (0, 2.5, 5, 12, 24 pmoles) (**Figure 13A**). En absence de protéines rajoutées, le site 5' proximal est davantage

utilisé (comparer pistes #2 et #3). En ajoutant A1 wt, le site d'épissage 5' distal est stimulé (pistes #4 et #5). Suite à la quantification des produits de l'épissage alternatif, l'activité de chaque mutant A1 est représentée graphiquement en fonction de sa concentration (**Figure 13B**). L'activité est calculée selon le rapport du compte radioactif de la bande correspondant aux lassos distaux sur le compte total des lassos distaux et proximaux. Les réactions d'épissage *in vitro* ont été répétées à trois reprises, tout comme la détection des comptes radioactifs de chaque réaction d'épissage. D'après le graphique, la courbe du mutant His-A1 23.01 est pratiquement identique à la courbe His-A1 wt, alors que la courbe du mutant 24.10 suggère une activité moyenne. Cependant, le mutant 210.11 n'a aucune activité jusqu'à 24 pmoles.

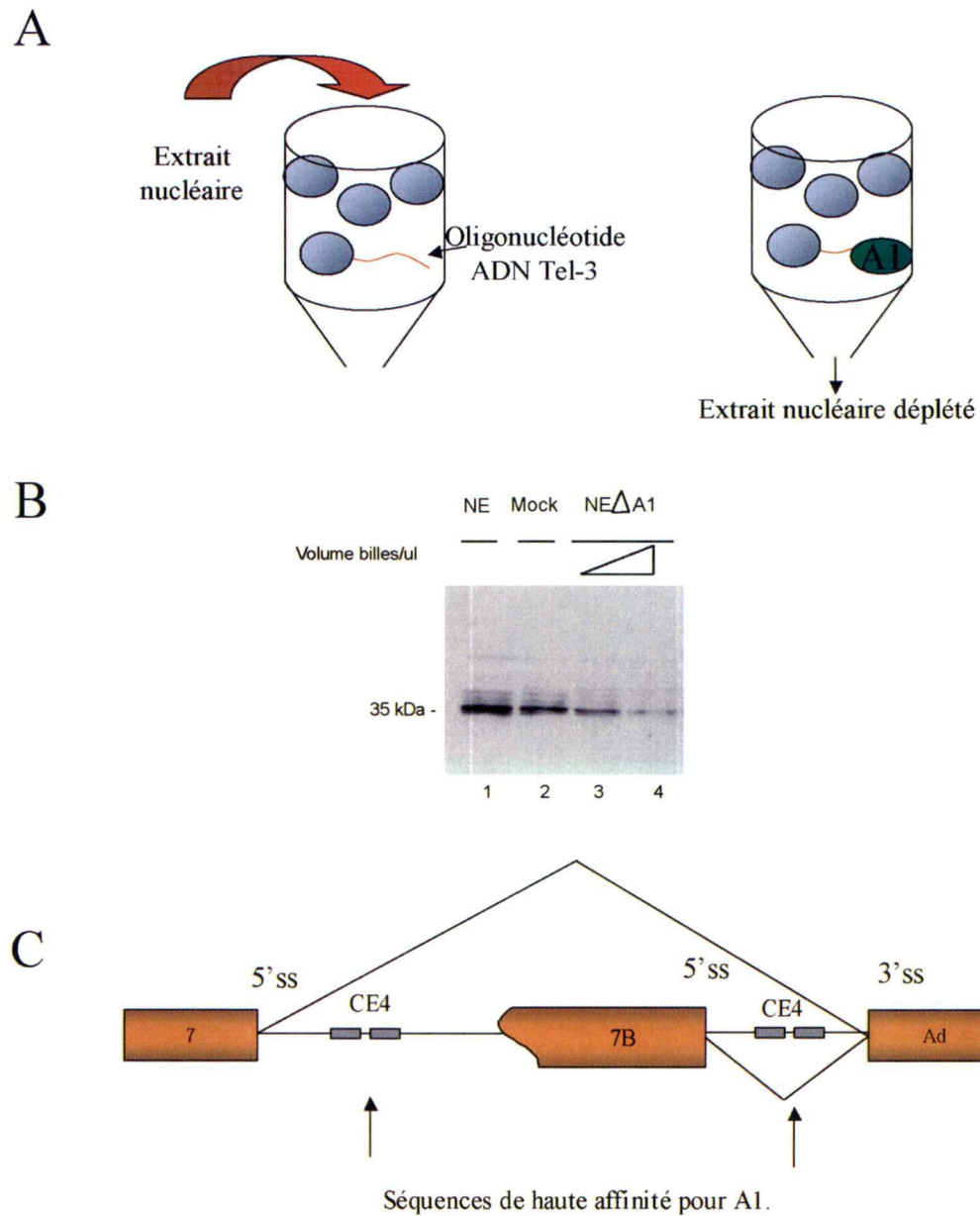
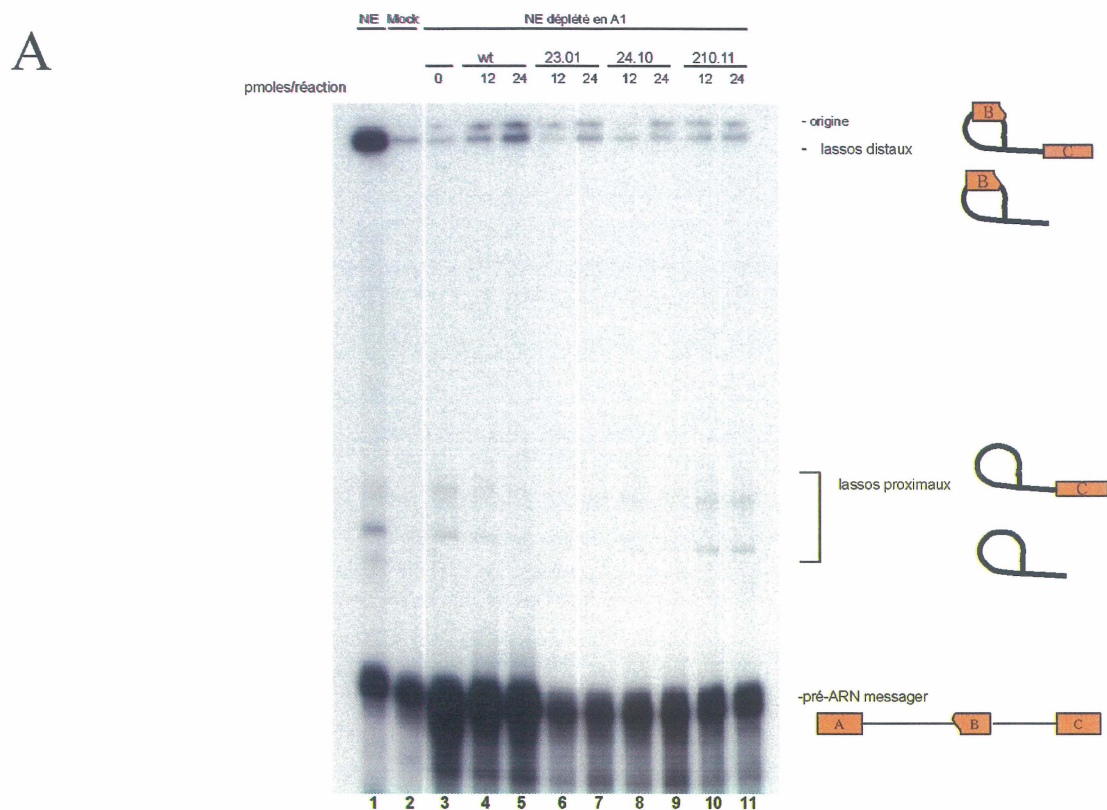


Figure 12. **Transcrit et déplétion d'extrait nucléaire.** A. Schéma représentant la déplétion d'un extrait cellulaire. Des oligonucléotides d'ADN contenant 3 répétitions télomériques simple brin pour A1 sont couplés à des billes d'agarose. Un extrait nucléaire de cellules HeLa est incubé sur la colonne pendant 20 min à 30 °C. B. Immunoblot représentant l'efficacité de la déplétion de hnRNP A1 endogène en fonction d'une quantité d'oligonucléotides croissante, à l'aide de l'anticorps anti-A1. C. Représentation schématique du transcrit dérivé du gène hnRNP A1. L'élément CE4 est composé de 2 sites de haute affinité pour A1.



B

Pourcentage de l'activité relative de hnRNP A1 dans l'épissage en fonction de la concentration de hnRNP A1 recombinantes.

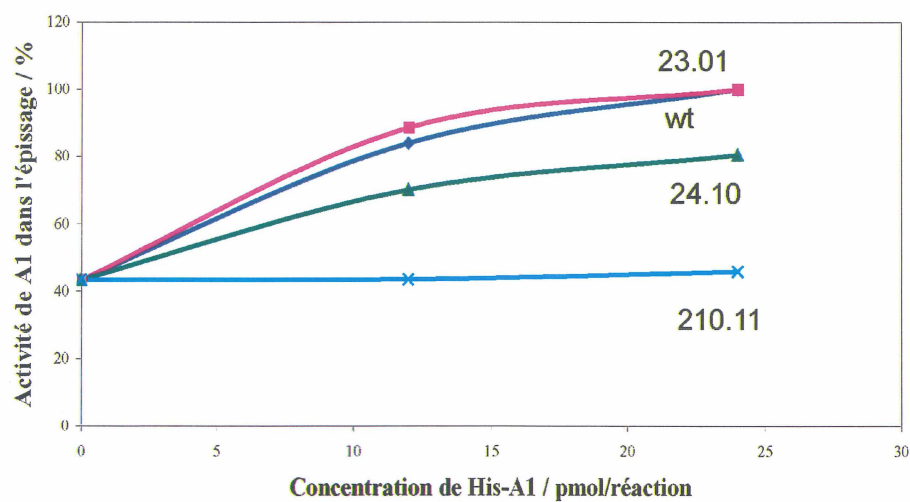


Figure 13. Certaines mutations dans le domaine riche en glycines affectent la sélection des sites d'épissage 5'. L'ARN précurseur C5' 4/4 a été incubé 2h à 30°C dans un extrait nucléaire déplété en présence de concentrations croissantes (0, 2.5, 5, 12, 24 pmoles) de différentes protéines His-A1 mutantes. Les réactions ont été déposées sur un gel polyacrylamide 11% contenant 8 M d'urée.

	# mutant	Acides aminés mutés	Liaison à l'ADN	Interaction A1-A1		Épissage alternatif
				Ade	lacZ	
1	21.01	Mutations silencieuses (3)	+++	+++	+++**	+++
2	20.04*	Glu -> Lys (126) Tyr -> Asn (128) Ser -> Cys (231)	N/D	+++	+	N/D
3	28.19*	Glu -> Gly(176) Ser -> Gly (252)	N/D	++	++	N/D
4	26.20*	Phe -> Tyr (153) Glu -> Gly (176) Gly -> Arg (258)	-	N/D	+++**	N/D
5	22.07	Asp -> Val (157) Iso -> Thr (253) Gln -> stop (302)	+	-	-	-
6	22.08*	Lys -> Mét (161) Changement de cadre de lecture, 6 aa différents, stop. (213)	+	-	-	N/D
7	21.24*	Gly -> Ser(229)	++	+++	++	-
8	21.15*	Gly -> Arg (219) Gly -> Arg (277)	++	N/D	+++**	-
9	21.07*	Glu -> Val (132)	++	-	-	-
10	21.14	Arg -> Leu (196) Phe -> Tyr (273) Asn -> Ser (301)	+++	+++	++	+++
11	23.01	Tyr -> Stop (305)	+++	++	++	+++
12	24.10	Asn -> Asp (215)	+++	++	++	++
13	210.11	Arg -> Cys (195) Ser -> Gly (197) Gly -> Ser (205) Gly -> Cys (291) Gln -> His (302)	+++	+	++	-

* Expression difficile sous forme de protéines recombinantes avec l'étiquette 6x histidines.

** Interaction positive, mais non quantifiée.

N/D : Donnée non disponible

Figure 14. **Tableau récapitulatif des caractéristiques des mutants de A1.** La nature et la position de chaque mutation sont décrites pour chaque mutant de hnRNP A1 (Voir la figure 8 pour une représentation graphique des protéines mutées). La quantification relative au niveau de l'interaction aux acides nucléiques, de l'interaction protéine-protéine par double hybride chez la levure sur milieu restrictif ou par un test b-galactosidase quantifié ou non, et l'activité dans la sélection des sites d'épissage 5' sont également inscrites.

DISCUSSION

Validité du système de répression traductionnelle

Pour déterminer les paramètres de départ du système de répression traductionnelle nécessaire pour évaluer la contribution de la liaison coopérative sur l'activité de A1, nous nous sommes basés sur notre connaissance de l'organisation des sites de liaison de A1 dans un gène dont l'épissage alternatif est contrôlé par A1. Sur l'ARN pré-messager de A1, on retrouve 4 sites de haute affinité pour hnRNP A1 de part et d'autre de l'exon alternatif 7B ; CE1a-CE1d en amont de 7B et CE4p-CE4m en aval (Hutchison et al., 2002). Sans être configurés comme la séquence optimale de A1 déterminée par SELEX, les éléments CE1 et CE4 ont des sites pour A1 très rapprochés. Chaque site est capable de lier hnRNP A1 indépendamment.

L'utilisation d'un système procaryote à la place d'un système eucaryote peut présenter un désavantage. Un système eucaryote permet des modifications post-traductionnelles sur les protéines pouvant occasionner une interaction protéine-protéine différente, causée par la phosphorylation par exemple. Ainsi, une mutation repêchée par le système de répression traductionnelle pourrait présenter une activité différente lorsque placée dans un contexte eucaryotique. L'effet de la mutation pourrait être renforcé ou diminué. Il est probable que des a.a. importants pour l'interaction A1-A1 ne soient pas détectés. La technique de répression traductionnelle a déjà été utilisée chez des cellules eucaryotes (Stripecke et al., 1994).

Un faux-positif ne comptant que des mutations silencieuses a été repêché (21.01) (Figure 14, ligne 1). Bien que nous n'ayons pas d'explications claires pour expliquer l'existence de ce faux-positif, il est possible que le séquençage n'est pas révélé la nature de la mutation. Comme l'expression de mutants de hnRNP A1 dans un système bactérien est parfois difficile, une forme dégradée d'une protéine normale pourrait parfois être responsable d'un phénotype mutant. Une autre possibilité serait que la nature des mutations causerait une liaison à un facteur nucléaire ou formerait une structure secondaire qui ralentirait la production de protéines. La limite de détection de l'activité de la β -galactosidase lors de la répression traductionnelle pourrait être atteinte. L'apparition de colonies bleues pâles, pour le mutant 21.01, favoriserait cette hypothèse.

L'interaction A1-A1 nécessite une liaison à une séquence d'acide nucléique simple brin

Des résultats non-publiés obtenus par Guillaume Lettre et Marco Blanchette suggèrent que la liaison de A1 à l'ARN permet une interaction A1-A1 plus efficace. Nos données tendent également en ce sens. Les mutants (22.07, 22.08, 21.24 et 21.07) ayant une déficience au niveau de l'interaction à l'ARN, montrent également une activité réduite à une interaction protéine-protéine suite à la quantification de l'activité de la β -galactosidase. Nous ne pouvons conclure pour les mutants 26.20 et

21.15 puisque leur force d'interaction protéine-protéine n'a pas été quantifiée. Nous savons seulement qu'ils ont une interaction positive. Il a été observé que la liaison de UP1 aux acides nucléiques induit un changement de conformation d'UP1 (Ding et al., 1999). Ce changement de conformation chez A1 permet peut-être d'exposer davantage la région nécessaire à l'interaction protéine-protéine. D'autres mutants devront être criblés afin de vérifier si une diminution de l'activité de liaison de A1 à l'ARN cause toujours une réduction de la force d'interaction protéine-protéine.

Le rôle du RRM2

Plusieurs mutants ayant une mutation dans le RRM2 montrent une déficience dans l'interaction A1-A1, mais la plupart ont d'autres mutations dans la protéine pouvant jouer un rôle. Seul le mutant 21.07 a perdu totalement sa capacité d'interaction et ne contient qu'une seule mutation dans le RRM2 (**Figure 14, ligne 9**). Ce mutant est caractérisé par le changement à la position 132 du RRM2, d'une glutamine en valine, c'est-à-dire d'un acide aminé avec une chaîne polaire hydrophile pour un a.a. avec une chaîne polaire hydrophobe. Son activité dans l'épissage est réduite malgré une assez bonne activité de liaison à l'ARN. Puisque son activité de liaison à l'ARN est toutefois un peu réduite, nous ne pouvons conclure que cet acide aminé est nécessaire à l'interaction protéine-protéine, car la réduction de l'activité dans l'épissage alternatif pourrait être causée par cette diminution dans la liaison à l'ARN. Selon notre modèle, l'interaction A1-A1 est favorisée lorsque A1 lie l'ARN d'abord. Un

changement de conformation permettant d'exposer les régions nécessaires à l'interaction A1-A1 suivrait. Chez la levure, A1 nécessite possiblement cette interaction à l'ARN avant d'activer le gène rapporteur. Donc, un défaut dans la liaison à l'ARN pourrait engendrer la perte d'interaction A1-A1. Toutes les protéines ayant une mutation dans le RRM2 semblent avoir une activité réduite au niveau de la liaison aux acides nucléiques et la plupart sont difficilement exprimées sous forme recombinante dans le système bactérien. Lorsque mesuré par un essai de liaison sur filtre en absence du domaine riche en glycines, le domaine RRM2 possède une activité de liaison à l'ADN télomérique plus faible par rapport au RRM1 d'approximativement de l'ordre de 3 fois (Dallaire et al., 2000). La contribution du RRM2 est peut-être plus importante pour l'ARN lorsque le domaine riche en glycines est présent et permettrait de stabiliser la liaison dans des conditions plus sévères de gel de retardement.

Il a déjà été montré, par la technique de double-hybride, que le domaine riche en glycines seul peut interagir avec lui-même (C. LeBel, résultats non-publiés) et que le domaine riche en glycines de A1 peut interagir avec le domaine riche en glycines de A2 (Cartegni et al., 1996). Cependant, les interactions A1-A1 et A1-A2, sous forme de protéines complètes, sont trois fois plus importantes que les interactions entre les domaines riche en glycines. Ceci suggère que d'autres régions dans le RRM2, ou ailleurs dans les 196 premiers a.a., sont nécessaires pour une interaction A1-A1 optimale.

Le domaine riche en glycines influence l'interaction protéine-protéine

Nous avons souvent plus d'une mutation par protéine, soit en moyenne trois à cinq. On ne sait donc pas toujours si le changement phénotypique est causé par une seule mutation ou par la combinaison de plusieurs mutations. Le nombre de mutations par protéine a été plus élevé que prévu. Il est possible qu'une erreur dans le changement des conditions de la *Taq* ce soit glissée lors de la mutagenèse, mais il semble davantage que ce soit le système de répression qui a sélectionné ces mutants. D'abord, il est probable qu'il faut affecter plus d'un acide aminé pour perturber suffisamment l'interaction protéine-protéine pour ensuite détecter cette perturbation par le système de répression traductionnelle. Mathématiquement, les chances sont plus faibles. Théoriquement, la *Taq* polymérase effectue normalement une mutation au 9000 nucléotides. Elle peut aussi causer l'insertion d'un nucléotide à tous les 41000 nucléotides et ainsi provoquer un changement de cadre de lecture. C'est un événement 5 fois plus rare et qui s'est produit pour le mutant 22.08. Si un tel événement s'est produit, il est très plausible de retrouver des protéines contenant 2 ou 3 mutations. D'autre part, environ 25% des mutants se terminent prématurément par un codon de terminaison. Cela constitue un événement fréquent lorsque statistiquement, il n'y a que 5% des chances de convertir un a.a. en codon stop. Pour le phénotype d'une perturbation de l'interaction protéine-protéine, une troncation du domaine riche en glycine était attendue puisque UP1 ne possède pas ce domaine.

Certaines protéines ont des mutations uniquement dans le domaine riche en glycines, et la technique de double-hybride chez la levure nous a permis de confirmer que tous les mutants ont une activité de dimérisation réduite par rapport à la protéine de type sauvage. Le mutant 21.14 contient trois mutations réparties dans le domaine riche en glycines. Le mutant 23.01 est raccourci de 15 a.a. à partir de l'extrémité carboxy-terminale. Ceci suggère que certains a.a. se retrouvant dans la queue terminale joueraient un rôle important dans l'activité de dimérisation de A1. Le mutant 24.10 a subi un changement d'un a.a. contenant une chaîne latérale hydrophile par un a.a. ayant une chaîne latérale chargée négativement, dans la boîte RGG. Cette charge négative apparaît dans une séquence chargée positivement causée par la présence de plusieurs arginines composant les répétitions RGG. Il est possible que cette charge négative perturbe la charge globale de cette région protéique participant possiblement à l'interaction protéine-protéine. Enfin, le mutant 210.11 contient 5 mutations dont 3 sont régionalisées dans la jonction des domaines RRM2 et riche en glycines. Le nombre de mutations est suffisamment important pour croire que la conformation structurale pourrait être fortement modifiée. Ultérieurement, ces mutants devront être décomposés afin de pouvoir analyser individuellement la contribution de chaque mutation. La diversité des mutations sur le plan de la localisation suggère qu'il n'y a pas une région critique à l'interaction protéine-protéine dans le domaine riche en glycines, mais que plusieurs régions du domaine pourraient contribuer à l'interaction A1-A1.

Rôle du domaine riche en glycines dans l'épissage

La protéine UP1, dont le domaine riche en glycines est absent, est incapable de promouvoir la sélection du site 5' distal sur un transcrit où deux sites d'épissage 5' sont en compétition. La déficience des mutants au niveau de l'interaction A1-A1 devrait donc aussi refléter leur activité dans l'épissage alternatif. Parmi les quatre mutants qui lient très bien les acides nucléiques, les mutants 21.14 et 23.01 ont été capables de favoriser l'activation du site 5' distal aussi efficacement que la protéine de type sauvage. Le mutant 24.10 a montré une activité réduite dans sa capacité à promouvoir l'utilisation du site 5' distal, alors qu'aucune activité n'a été observée pour le mutant 210.11. Donc, il n'y a pas de corrélation entre l'activité des mutants A1 dans l'épissage et la force d'interaction mesurée en double-hybride. Afin d'expliquer les résultats avec les mutants 21.14 et 23.01, nous pouvons suggérer que l'épissage *in vitro* par addition de protéines A1 recombinantes dans un extrait déplété, n'est peut-être pas aussi sensible aux mutations de A1 que les tests de répression traductionnelle et de double-hybride. Malgré la déplétion, une quantité de protéines endogènes hnRNP A1/A1B, A2/B1 subsiste toujours dans l'extrait nucléaire de cellules HeLa. L'épissage *in vitro* dans un extrait déplété étant moins efficace que dans un extrait normal, cette quantité de protéines endogènes diminue aussi la sensibilité de l'essai. Ces facteurs empêchent peut-être l'observation d'un lien clair entre la force d'interaction A1-A1 et le déplacement vers le site d'épissage 5' distal.

Quatre sites de liaison de haute affinité pour A1, similaire à la configuration de la séquence optimale déterminée par SELEX, ont été retrouvés sur chaque intron flanquant l'exon alternatif 7B de l'ARN pré-messager de A1 (Hutchison et al., 2002). Cependant, nous ne savons pas si 2 sites optimaux pour A1 requièrent la présence de 2 ou de 4 molécules A1. La répression du système de répression traductionnelle par A1 et l'activité de modulation de l'épissage alternatif *in vitro* par A1, sont peut-être facilitées par la contribution de 4 molécules A1. Par contre, le double-hybride chez la levure mesure la force d'interaction seulement entre deux molécules. Dans un contexte de 4 molécules, le domaine riche en glycines pourrait participer à l'interaction protéine-protéine en présentant son domaine différemment. La configuration adoptée par les 4 molécules pourrait expliquer pourquoi il ne semble pas exister de corrélation dans nos résultats, entre l'interaction A1-A1 et l'activité de A1 dans l'épissage.

HnRNP A1 modélisée dans l'épissage alternatif

En conclusion, l'information rapportée par les différents mutants A1 repêchés par la technique de répression traductionnelle reste en accord avec le modèle proposé sur le mécanisme de A1 dans l'épissage alternatif. Il suggère qu'A1 favoriserait d'abord une liaison à l'ARN pour ensuite procéder à une interaction A1-A1. Il nous indique aussi que le RRM2 jouerait un rôle important sur l'activité de liaison aux acides nucléiques, mais aussi qu'il pourrait possiblement intervenir dans l'activité d'interaction protéine-

protéine. L'interaction A1-A1 serait possible par l'entremise de plusieurs régions dans le domaine riche en glycines, et probablement dans le RRM2.

REMERCIEMENTS

J'aimerais d'abord remercier tout spécialement mon directeur de recherche, le Dr. Benoit Chabot, pour sa sagesse et sa patience au cours de mon cheminement pendant ma maîtrise. Ses précieux conseils et sa disponibilité surprenante m'ont permis de découvrir le vrai visage de la recherche sous des traits qui m'étaient encore inconnus. Sa facilité à lever les obstacles et son esprit critique resteront un exemple à suivre. Je remercie la Faculté de Médecine de m'avoir offert une bourse. Je remercie spécialement tous les membres du laboratoire pour l'ambiance positive qui émanait chaque jour dans notre travail, tout comme les loisirs avec qui j'ai partagé dans un contexte un peu moins scientifique. Je souligne plus spécialement Stephen Hutchison et Martin Simard pour les nombreuses discussions qui ont apporté énormément à l'avancement de la science et à d'autres domaines plus ésotériques. Merci à Lucie Bolduc et Mélanie Cordeau pour ma formation et leur bonne humeur. Merci également à Jonathan Villemaire, Daniel Garneau, Timothée Revil, Jean-François Fisette, Catherine LeBel et Faiz Ul Nasim pour le bon temps passé à se côtoyer entre les centrifugeuses et pour leurs conseils judicieux. Un grand merci à Maryse, Annie et Aline qui m'ont inculqué les bonnes méthodes de laboratoire. Enfin, je tiens à remercier ma famille, mes ami(e)s et Isabelle pour leur support moral.

Martin Simard, Catherine LeBel, Daniel Garneau, Timothée Revil et Aline Simoneau ont contribué activement à ce projet.

BIBLIOGRAPHIE

1. Adam, S. A., Nakagawa, T., Swanson, M. S., Woodruff, T. K., et Dreyfuss, G. (1986). mRNA polyadenylate-binding protein: gene isolation and sequencing and identification of a ribonucleoprotein consensus sequence. *Mol Cell Biol* 6, 2932-2943.
2. Ausubel, F. R., ed. (1998). *Current Protocols in Molecular Biology* (John Wiley & Sons, Inc.).
3. Bai, Y., Lee, D., Yu, T., et Chasin, L. A. (1999). Control of 3' splice site choice in vivo by ASF/SF2 and hnRNP A1. *Nucleic Acids Res* 27, 1126-1134.
4. Ben-David, Y., Bani, M. R., Chabot, B., De Koven, A., et Bernstein, A. (1992). Retroviral insertions downstream of the heterogeneous nuclear ribonucleoprotein A1 gene in erythroleukemia cells: evidence that A1 is not essential for cell growth. *Mol Cell Biol*, 4449-4455.
5. Berget, S. M., Moore, C., et Sharp, P. A. (1977). Spliced segments at the 5' terminus of adenovirus 2 late mRNA. *Proc Natl Acad Sci U S A* 74, 3171-3175.
6. Biamonti, G., et Riva, S. (1994). New insights into the auxiliary domains of eukaryotic RNA binding proteins. *FEBS Lett* 340, 1-8.
7. Black, D. L. (2003). Mechanisms of alternative pre-messenger RNA splicing. *Annu Rev Biochem*, 72 :291-336.

8. Blanchette, M., et Chabot, B. (1997). A highly stable duplex structure sequesters the 5' splice site region of hnRNP A1 alternative exon 7B. *RNA* 3, 405-419.
9. Blanchette, M., et Chabot, B. (1999). Modulation of exon skipping by high-affinity hnRNP A1-binding sites and by intron elements that repress splice site utilization. *EMBO J* 18, 1939-1952.
10. Bolduc, L., Labrecque, B., Cordeau, M., Blanchette, M., et Chabot, B. (2001). Dimethyl sulfoxide affects the selection of splice sites. *J Biol Chem* 276, 17597-17602.
11. Brow, D. A. (2002). Allosteric cascade of spliceosome activation. *Annu Rev Genet* 36, 333-360.
12. Burd, C. G., et Dreyfuss, G. (1994). RNA binding specificity of hnRNP A1: significance of hnRNP A1 high-affinity binding sites in pre-mRNA splicing. *EMBO J* 13, 1197-1204.
13. Caceres, J. F., et Kornblihtt, A. R. (2002). Alternative splicing: multiple control mechanisms and involvement in human disease. *Trends Genet* 18, 186-193.
14. Caputi, M., Mayeda, A., Krainer, A. R., et Zahler, A. M. (1999). hnRNP A/B proteins are required for inhibition of HIV-1 pre-mRNA splicing. *EMBO J* 18, 4060-4067.
15. Caputi, M., et Zahler, A. M. (2001). Determination of the RNA binding specificity of the heterogeneous nuclear ribonucleoprotein (hnRNP) H/H'/F/2H9 family. *J Biol Chem* 276, 43850-43859.

16. Caputi, M., et Zahler, A.M. (2002) SR proteins and hnRNP H regulate the splicing of the HIV-1 *tev*-specific exon 6D. *EMBO J.* 20, 845-855.
17. Cartegni, L., Maconi, M., Morandi, E., Cobianchi, F., Riva, S., et Biamonti, G. (1996). hnRNP A1 selectively interacts through its Gly-rich domain with different RNA-binding proteins. *J Mol Biol* 259, 337-348.
18. Celander, D. W., Abelson, J.N., ed. (2000). RNA-Ligand Interactions Part B (Academic Press).
19. Chabot, B. (1996). Directing alternative splicing: cast and scenarios. *Trends Genet* 12, 472-478.
20. Chabot, B., Blanchette, M., Lapierre, I., et La Branche, H. (1997). An intron element modulating 5' splice site selection in the hnRNP A1 pre-mRNA interacts with hnRNP A1. *Mol Cell Biol* 17, 1776-1786.
21. Chabot, B., LeBel, C., Hutchison, S., Nasim, F. U., et Simard, M. J. (2003). Heterogeneous nuclear ribonucleoprotein particle A/B proteins and the control of alternative splicing of the mammalian heterogeneous nuclear ribonucleoprotein particle A1 pre-mRNA. In *Progress in Molecular and Subcellular Biology*, P. Jeanteur, ed.
22. Chen, C.D., Kobayashi, R., Helfman, D.M., (1999). Binding of hnRNP H to an exonic splicing silencer is involved in the regulation of alternative splicing of the rat β -tropomyosin gene. *Genes and Dev.* 13, 593-606.
23. Choi, Y. D., et Dreyfuss, G. (1984). Isolation of the heterogeneous nuclear RNA-ribonucleoprotein complex (hnRNP): a unique supramolecular assembly. *Proc Natl Acad Sci U S A* 81, 7471-7475.

24. Chow, L. T., Gelinas, R. E., Broker, T. R., et Roberts, R. J. (1977). An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA. *Cell* 12, 1-8.
25. Cobiauchi, F., Karpel, R. L., Williams, K. R., Notario, V., et Wilson, S. H. (1988). Mammalian heterogeneous nuclear ribonucleoprotein complex protein A1. Large-scale overproduction in *Escherichia coli* and cooperative binding to single-stranded nucleic acids. *J Biol Chem* 263, 1063-1071.
26. Dallaire, F., Dupuis, S., Fiset, S., et Chabot, B. (2000). Heterogeneous nuclear ribonucleoprotein A1 and UP1 protect mammalian telomeric repeats and modulate telomere replication in vitro. *J Biol Chem* 275, 14509-14516.
27. Damgaard, C. K., Tange, T. O., et Kjems, J. (2002). hnRNP A1 controls HIV-1 mRNA splicing through cooperative binding to intron and exon splicing silencers in the context of a conserved secondary structure. *RNA* 8, 1401-1415.
28. Dignam, J. D., Lebovitz, R. M., et Roeder, R. G. (1983). Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res* 11, 1475-1489.
29. Ding, J., Hayashi, M. K., Zhang, Y., Manche, L., Krainer, A. R., et Xu, R. M. (1999). Crystal structure of the two-RRM domain of hnRNP A1 (UP1) complexed with single-stranded telomeric DNA. *Genes Dev* 13, 1102-1115.
30. Dreyfuss, G., Swanson, M. S., et Pinol-Roma, S. (1988). Heterogeneous nuclear ribonucleoprotein particles and the pathway of mRNA formation. *Trends Biochem Sci* 13, 86-91.

31. Graveley, B. R. (2000). Sorting out the complexity of SR protein functions. *RNA* 6, 1197-1211.
32. Hamilton, B. J., Burns, C. M., Nichols, R. C., et Rigby, W. F. (1997). Modulation of AUUUA response element binding by heterogeneous nuclear ribonucleoprotein A1 in human T lymphocytes. The roles of cytoplasmic location, transcription, and phosphorylation. *J Biol Chem* 272, 28732-28741.
33. Hay, D. C., Kemp, G. D., Dargemont, C., et Hay, R. T. (2001). Interaction between hnRNPA1 and IkappaBalpha is required for maximal activation of NF-kappaB-dependent transcription. *Mol Cell Biol* 21, 3482-3490.
34. Hirose, Y., et Manley, J. L. (2000). RNA polymerase II and the integration of nuclear events. *Genes Dev* 14, 1415-1429.
35. Hou, V. C., Lersch, R., Gee, S. L., Ponthier, J. L., Lo, A. J., Wu, M., Turck, C. W., Koury, M., Krainer, A. R., Mayeda, A., et Conboy, J. G. (2002). Decrease in hnRNP A/B expression during erythropoiesis mediates a pre-mRNA splicing switch. *EMBO J* 21, 6195-6204.
36. Hutchison, S., LeBel, C., Blanchette, M., et Chabot, B. (2002). Distinct sets of adjacent heterogeneous nuclear ribonucleoprotein (hnRNP) A1/A2 binding sites control 5' splice site selection in the hnRNP A1 mRNA precursor. *J Biol Chem* 277, 29745-29752.
37. Innis, ed. (1990). *PCR Protocols; A guide to methods and applications*, Academic Press edn (Harcourt Brace Jovanovich, publishers).

38. Izaurralde, E., Jarmolowski, A., Beisel, C., Mattaj, I. W., Dreyfuss, G., et Fischer, U. (1997). A role for the M9 transport signal of hnRNP A1 in mRNA nuclear export. *J Cell Biol* 137, 27-35.
39. Jain, C., et Belasco, J. G. (1996). A structural model for the HIV-1 Rev-RRE complex deduced from altered-specificity rev variants isolated by a rapid genetic strategy. *Cell* 87, 115-125.
40. Jain, C., et Belasco, J. G. (2001) Structural model for the cooperative assembly of HIV-1 Rev multimers on the RRE as deduced from analysis of assembly-defective mutants. *Mol. Cell.* 7, 603-614.
41. James, P., Halladay, J., et Craig, E. A. (1996). Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast. *Genetics* 144, 1425-1436.
42. Kamma, H., Portman, D. S., et Dreyfuss, G. (1995). Cell type-specific expression of hnRNP proteins. *Exp Cell Res* 221, 187-196.
43. Kay, B. K., Sawhney, R. K., et Wilson, S. H. (1990). Potential for two isoforms of the A1 ribonucleoprotein in *Xenopus laevis*. *Proc Natl Acad Sci U S A* 87, 1367-1371.
44. Kiledjian, M., et Dreyfuss, G. (1992). Primary structure and binding activity of the hnRNP U protein: binding RNA through RGG box. *EMBO J* 11, 2655-2664.
45. Kim, S., Merrill, B. M., Rajpurohit, R., Kumar, A., Stone, K. L., Papov, V. V., Schneiders, J. M., Szer, W., Wilson, S. H., Paik, W. K., et Williams, K. R. (1997). Identification of N(G)-methylarginine residues in human

- heterogeneous RNP protein A1: Phe/Gly-Gly-Gly-Arg-Gly-Gly-Gly/Phe is a preferred recognition motif. *Biochemistry* **36**, 5185-5192.
46. Kramer, A. (1996). The structure and function of proteins involved in mammalian pre-mRNA splicing. *Annu Rev Biochem* **65**, 367-409.
47. LaBranche, H., Dupuis, S., Ben-David, Y., Bani, M. R., Wellinger, R. J., et Chabot, B. (1998). Telomere elongation by hnRNP A1 and a derivative that interacts with telomeric repeats and telomerase. *Nat Genet* **19**, 199-202.
48. Lavigueur, A., La Branche, H., Kornblihtt, A. R., et Chabot, B. (1993). A splicing enhancer in the human fibronectin alternate ED1 exon interacts with SR proteins and stimulates U2 snRNP binding. *Genes Dev* **7**, 2405-2417.
49. Li, B., Wachtel, C., Miriami, E., Yahalom, G., Friedlander, G., Sharon, G., Sperling, R., et Sperling, J. (2002). Stop codons affect 5' splice site selection by surveillance of splicing. *Proc Natl Acad Sci U S A* **99**, 5277-5282.
50. Lou, H., Helfman, D. M., Gagel, R. F., et Berget, S. M. (1999). Polypyrimidine tract-binding protein positively regulates inclusion of an alternative 3'-terminal exon. *Mol Cell Biol* **19**, 78-85.
51. Maniatis, T., et Reed, R. (2002). An extensive network of coupling among gene expression machines. *Nature* **416**, 499-506.
52. Marchand, V., Mereau, A., Jacquenet, S., Thomas, D., Mougin, A., Gattoni, R., Stevenin, J., et Branlant, C. (2002). A Janus splicing regulatory element modulates HIV-1 tat and rev mRNA production by coordination of hnRNP A1 cooperative binding. *J Mol Biol* **323**, 629-652.

53. Mayeda, A., et Krainer, A. R. (1992). Regulation of alternative pre-mRNA splicing by hnRNP A1 and splicing factor SF2. *Cell* 68, 365-375.
54. Mayeda, A., Munroe, S. H., Cáceres, J. F., et Krainer, A. R. (1994). Function of conserved domains of hnRNP A1 and other hnRNP A/B proteins. *EMBO J* 13, 5483-5495.
55. Mayeda, A., Munroe, S. H., Xu, R. M., et Krainer, A. R. (1998). Distinct functions of the closely related tandem RNA-recognition motifs of hnRNP A1. *RNA* 4, 1111-1123.
56. McPherson, J. D., Marra, M., Hillier, L., Waterston, R. H., Chinwalla, A., Wallis, J., Sekhon, M., Wylie, K., Mardis, E. R., Wilson, R. K., *et al.* (2001). A physical map of the human genome. *Nature* 409, 934-941.
57. Michael, W. M., Choi, M., et Dreyfuss, G. (1995). A nuclear export signal in hnRNP A1: a signal-mediated, temperature-dependent nuclear protein export pathway. *Cell* 83, 415-422.
58. Min, H., Chan, R.C., Black D.L. (1995). The generally expressed hnRNP F is involved in a neural-specific pre-mRNA splicing event. *Genes and Dev.* 9, 2659-2671.
59. Misteli, T. (1999). RNA splicing: What has phosphorylation got to do with it? *Curr Biol* 9, R198-200.
60. Nasim, F. U., Hutchison, S., Cordeau, M., et Chabot, B. (2002). High-affinity hnRNP A1 binding sites and duplex-forming inverted repeats have similar effects on 5' splice site selection in support of a common looping out and repression mechanism. *RNA* 8, 1078-1089.

61. Nichols, R. C., Wang, X. W., Tang, J., Hamilton, B. J., High, F. A., Herschman, H. R., et Rigby, W. F. (2000). The RGG domain in hnRNP A2 affects subcellular localization. *Exp Cell Res* 256, 522-532.
62. Otake, L. R., Scamborova, P., Hashimoto, C., et Steitz, J. A. (2002). The divergent U12-type spliceosome is required for pre-mRNA splicing and is essential for development in *Drosophila*. *Mol Cell* 9, 439-446.
63. Peled-Zehavi, H., Berglund, J. A., Rosbash, M., et Frankel, A. D. (2001). Recognition of RNA branch point sequences by the KH domain of splicing factor 1 (mammalian branch point binding protein) in a splicing factor complex. *Mol Cell Biol* 21, 5232-5241.
64. Qiagen (2001). The QIAexpressionist: A handbook for high-level expression and purification of 6xHis-tagged proteins, 5e édition edn).
65. Sambrook, J., Fritsch, E. F., Maniatis, T. (1989). Molecular cloning. In a laboratory manual.
66. Schmucker, D., Clemens, J. C., Shu, H., Worby, C. A., Xiao, J., Muda, M., Dixon, J. E., et Zipursky, S. L. (2000). *Drosophila* Dscam is an axon guidance receptor exhibiting extraordinary molecular diversity. *Cell* 101, 671-684.
67. Shamoo, Y., Krueger, U., Rice, L. M., Williams, K. R., et Steitz, T. A. (1997). Crystal structure of the two RNA binding domains of human hnRNP A1 at 1.75 Å resolution. *Nat Struct Biol* 4, 215-222.
68. Simard, M. J., et Chabot, B. (2002). SRp30c is a repressor of 3' splice site utilization. *Mol Cell Biol* 22, 4001-4010.

69. Smith, C. W.J., Valcárcel, J., (2000). Alternative pre-mRNA splicing: the logic of combinatorial control. *TIBS* 25, 381-388.
70. Stevens, S. W., Ryan, D. E., Ge, H. Y., Moore, R. E., Young, M. K., Lee, T. D., et Abelson, J. (2002). Composition and functional characterization of the yeast spliceosomal penta-snRNP. *Mol Cell* 9, 31-44.
71. Stripecke, R., Oliveira, C.O., McCarthy, J.E.G., et Hentze, M.W., (1994). Proteins binding to 5' untranslated region sites : a general mechanism for translational regulation of mRNAs in human and yeast cells. *Mol. Cell. Biol.* 14, 5898-5909.
72. Tange, T. O., Damgaard, C. K., Guth, S., Valcarcel, J., et Kjems, J. (2001). The hnRNP A1 protein regulates HIV-1 tat splicing via a novel intron silencer element. *EMBO J* 20, 5748-5758.
73. Vitali, J., Ding, J., Jianzhong, J., Zhang, Y., Krainer, A.R., Xu, R.M., (2002) Correlated alternative side chain conformations in the RNA-recognition motif of heterogeneous nuclear ribonucleoprotein A1. *Nucleic Acids Res* 30, 1531-1538.
74. Wang, Z., Hoffmann, H. M., et Grabowski, P. J. (1995). Intrinsic U2AF binding is modulated by exon enhancer signals in parallel with changes in splicing activity. *RNA* 1, 21-35.
75. Will, C. L., et Luhrmann, R. (1997). Protein functions in pre-mRNA splicing. *Curr Opin Cell Biol* 9, 320-328.
76. Zahler, A. M., Lane, W. S., Stolk, J. A., et Roth, M. B. (1992). SR proteins: a conserved family of pre-mRNA splicing factors. *Genes Dev* 6, 837-847.

77. Zhu, J., Mayeda, A., et Krainer, A. R. (2001). Exon identity established through differential antagonism between exonic splicing silencer-bound hnRNP A1 and enhancer-bound SR proteins. *Mol Cell* 8, 1351-1361.

ANNEXE I

Bolduc, L., Labrecque, B., Cordeau, M., Blanchette, M., and Chabot, B. (2001). Dimethyl sulfoxide affects the selection of splice sites. *J Biol Chem* 276, 17597-17602.

J'ai contribué à l'article : "Dimethyl sulfoxide affects the selection of splice sites" en répétant l'expérience dont les résultats sont présentés à la figure 5 : "Me₂SO rescues splicing in HeLa S100 extract". J'ai également effectué les expériences dont les résultats sont présentés à la figure 6 : "Me₂SO activates SR proteins", et à la figure 7 : "Modulating and stimulating activities of DMF and formamide". Ceci comprend également la production de protéines recombinantes.

Dimethyl Sulfoxide Affects the Selection of Splice Sites*

Received for publication, December 28, 2000, and in revised form, February 19, 2001
Published, JBC Papers in Press, March 1, 2001, DOI 10.1074/jbc.M011769200

Lucie Bolduc‡, Benoit Labrecque, Mélanie Cordeau, Marco Blanchette§, and Benoit Chabot¶

From the Département de Microbiologie et d'Infectiologie, Faculté de Médecine, Université de Sherbrooke, Sherbrooke, Québec, Canada J1H 5N4

Depending on the cell lines and cell types, dimethyl sulfoxide (Me_2SO) can induce or block cell differentiation and apoptosis. Although Me_2SO treatment alters many levels of gene expression, the molecular processes that are directly affected by Me_2SO have not been clearly identified. Here, we report that Me_2SO affects splice site selection on model pre-mRNAs incubated in a nuclear extract prepared from HeLa cells. A shift toward the proximal pair of splice sites was observed on pre-mRNAs carrying competing 5'-splice sites or competing 3'-splice sites. Because the activity of recombinant hnRNP A1 protein was similar when added to extracts containing or lacking Me_2SO , the activity of endogenous A1 proteins is probably not affected by Me_2SO . Notably, in a manner reminiscent of SR proteins, Me_2SO activated splicing in a HeLa S100 extract. Moreover, the activity of recombinant SR proteins in splice site selection *in vitro* was improved by Me_2SO . Polar solvents like DMF and formamide similarly modulated splice site selection *in vitro* but formamide did not activate a HeLa S100 extract. We propose that Me_2SO improves ionic interactions between splicing factors that contain RS-domains. The direct impact of Me_2SO on alternative splicing may explain, at least in part, the different and sometimes opposite effects of Me_2SO on cell differentiation and apoptosis.

Me_2SO ¹ is a polar solvent used to promote cell differentiation of tumor cell lines. For example, the treatment of mouse erythroleukemic and neuroblastoma cells with 2% Me_2SO induces morphological changes and differentiation in red blood cells and neurons, respectively (e.g. see Refs. 1, 2). Me_2SO also induces differentiation of the human U937 monoblast leukemia cell line into monocyte/macrophage (3) and stimulates the differentiation of a human ovarian adenocarcinoma cell line (4). Paradoxically, Me_2SO prevents the terminal differentiation of myoblasts (5, 6), inhibits the differentiation of adipocytes (7),

blocks the differentiation of antibody-producing plasma cells (8), and interferes with the differentiation of chick embryo chondrocytes (9). Whereas Me_2SO has been used to induce apoptosis in some cell lines (10, 11), it inhibits cell density-dependent apoptosis of CHO cells (6). Thus, depending on the cell line, Me_2SO can have completely different effects on differentiation and apoptosis.

The cellular mechanisms that are affected by Me_2SO remain unclear. Because Me_2SO facilitates DNA uptake during transfection procedures (e.g. see Ref. 12), Me_2SO has been proposed to affect the integrity of cell membranes. Because Me_2SO alters protein kinase C activity and the expression of integrin complexes (6, 13), Me_2SO may alter intracellular signaling processes, which in turn may have a broad impact on many aspects of gene expression. Me_2SO treatment promotes changes in the abundance of certain mRNAs and in the ratio of spliced isoforms (14–17). Among the genes reported to be affected in their alternative splicing is the NCAM pre-mRNA. A 2% Me_2SO treatment of N2a cells promotes an increase in the inclusion of neuro-specific NCAM exon 18 (18, 19). Me_2SO alters the alternative splicing of other genes including the amyloid precursor protein (20), the serotonin 5-HT₃ receptor-A mRNA (21), and p53 (22, 23). Me_2SO has also been associated with an effect on c-Myc mRNA elongation, maturation, and stability (23–25), and on the translation of some mRNAs (1). Whether any of the above changes result from a direct effect of Me_2SO on RNA synthesis, maturation, and/or stability is currently unknown.

Because treating cells with Me_2SO can have a strong effect on the alternative splicing of many pre-mRNAs and because the mechanism of action of Me_2SO remains unclear, we performed a series of experiments in nuclear extracts to assess whether Me_2SO directly affects the activity of the splicing machinery. We find that Me_2SO can have drastic effects both on 5'-splice site and on 3'-splice site selection *in vitro*. Notably, other solvents of the same category (e.g. DMF and formamide) also perturb splice site selection.

EXPERIMENTAL PROCEDURES

Treatment of Cells with Me_2SO and *In Vivo* Alternative Splicing Assays— Me_2SO was purchased from various suppliers including EM Science and Fisher Scientific Inc. DMF and formamide were from Calbiochem. N2a and HeLa cells were cultured at 37 °C in Dulbecco's modified Eagle's medium supplemented with 10% bovine calf serum. For treatment with Me_2SO , medium containing 2% bovine calf serum was used. Following treatment, total RNA was isolated using the guanidinium-HCl protocol as described in Chabot (26). RNase T1 protection assay was performed according to Melton *et al.* (27) using a uniformly labeled 530-nt NCAM antisense RNA probe. Exon 17/exon 19 splicing yields a 303-nt protected fragment while the inclusion of exon 18 produces a 452-nt fragment. Products were resolved on a 5% denaturing acrylamide gel. The reverse transcriptase-PCR assay used to amplify products corresponding to exon 7B inclusion and exclusion has been described in Chabot *et al.* (28).

Substrate Pre-mRNAs and *In Vitro* Splicing Assays—pC5' -/-, pC5' 4/4 and pC3' -/- have been described in Blanchette and Chabot (29). pNCAM3' was constructed by substituting the 3'-splice site of exon 7B

* This work was supported in part by a grant from the National Cancer Institute of Canada with funds from the Canadian Cancer Society (to B. C.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ Recipient of a studentship from the Fonds pour la Formation de Chercheurs et l'Aide à la Recherche (FCAR).

§ Recipient of a studentship from the Medical Research Council of Canada. Present Address: Dept. of Molecular and Cell Biology, University of California, Berkeley, CA 94720-3204.

¶ Research Scholar from the Fond de la Recherche en Santé du Québec and member of the Sherbrooke RNA/RNP group supported by the FCAR. To whom correspondence should be addressed. Tel.: 819-564-5295; Fax: 819-564-5392; E-mail: b.chabot@courrier.usherb.ca.

¹ The abbreviations used are: Me_2SO , dimethyl sulfoxide; DMF, *N,N*-dimethylformamide; nt, nucleotide(s); bp, base pairs; hnRNP, heterogeneous nuclear ribonucleoprotein; NCAM, neural cell adhesion molecule; GST, glutathione *S*-transferase; PCR, polymerase chain reaction.

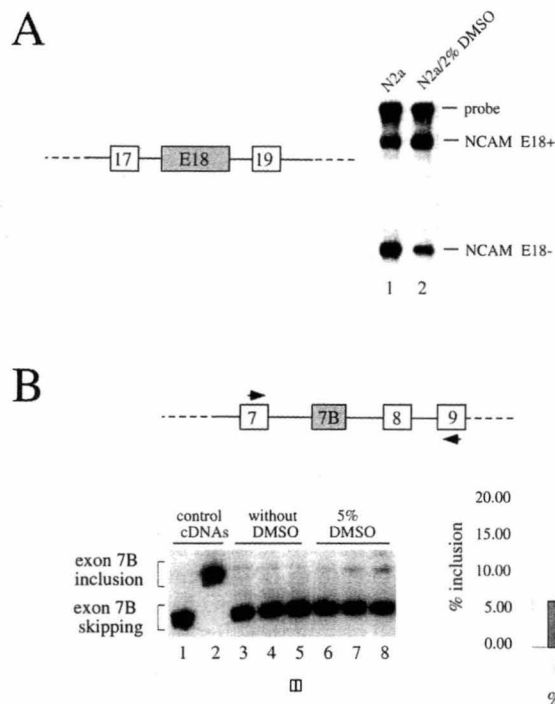


FIG. 1. Me_2SO promotes exon inclusion *in vivo*. **A**, mouse N2a cells were treated 48 h with 2% Me_2SO . Total RNA was analyzed for changes in the alternative splicing of NCAM exon 18. A RNase T1 protection assay was used to monitor the ratio of exon 18 inclusion (E18+) or exclusion (E18-). **B**, HeLa cells were treated with 5% Me_2SO for 5 h. Three dishes of cells were tested for each treatment and total RNA was analyzed for changes in the alternative splicing of exon 7B in the hnRNP A1 pre-mRNA. The percentage of exon 7B inclusion on endogenous A1 transcripts was determined by using a reverse transcriptase-PCR assay as described in Chabot *et al.* (28). Control PCR reactions were performed with plasmids containing the cDNA from A1 (lacking exon 7B, lane 1), or A1^B (containing exon 7B, lane 2). The values were plotted as percentage of inclusion on a histogram that shows standard deviations.

and exon 7B sequences in pC3' -/- for the equivalent regions of alternative exon 18 of the mouse NCAM gene (403 bp of exon 18 and 111 bp of upstream intron sequences). Splicing substrates were produced from plasmids linearized with *ScaI*, and transcribed with T3 RNA polymerase in the presence of cap analogue and [α - ^{32}P]UTP (Amersham Pharmacia Biotech). RNA purification was performed as described in Chabot (26). HeLa nuclear extracts and S100 extracts were prepared (30) and used in splicing reactions as described previously (28). Although Me_2SO , DMF, and formamide were always added last, the order of addition did not affect the outcome.

Recombinant A1 and SR Proteins—Recombinant GST-A1, GST-SRp30c were purified using a glutathione-Sepharose column (Amersham Pharmacia Biotech). Bacterial lysis was in buffer A (50 mM piperazine-HCl, pH 9.8, 0.5 M NaCl, 1 mM EDTA, 1 mM dithiothreitol, 20 $\mu\text{g}/\text{ml}$ bacitracin, 1 mM benzamidin, 0.5 mM phenylmethylsulfonyl fluoride in the presence of 3 mg/ml lysozyme and 1% Triton X-100. Elution from the columns was performed with buffer B (200 mM piperazine-HCl, pH 9.8, 0.5 M NaCl, 1 mM EDTA, 1 mM dithiothreitol, 20 mM reduced glutathione). The purified proteins were dialysed against buffer D (20 mM Hepes pH 7.9, 100 mM KCl, 20% glycerol, 1 mM EDTA, 0.5 mM dithiothreitol, 0.5 mM phenylmethylsulfonyl fluoride).

RESULTS

Me_2SO Affects Alternative Pre-mRNA Splicing *in Vivo*— Me_2SO can promote cell differentiation, a process that is often associated with a change in the alternative splicing profile of specific genes. One example of this effect is found in the mouse N2a neuroblastoma cell line. The treatment of N2a cells with 2% Me_2SO induces neuronal cell differentiation and improves the frequency of inclusion of the neurospecific exon 18 in the NCAM pre-mRNA (Refs. 2, 31; Fig. 1A). A similar effect was observed on the hnRNP A1 pre-mRNA. In this case, we moni-

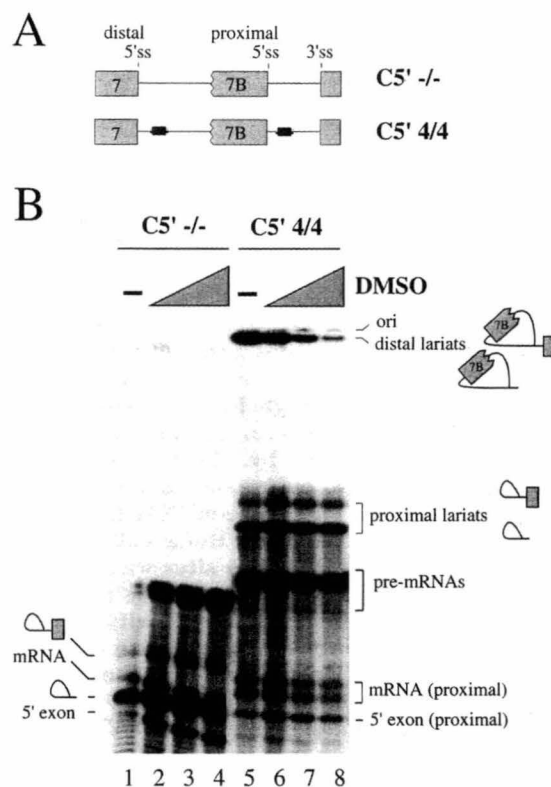


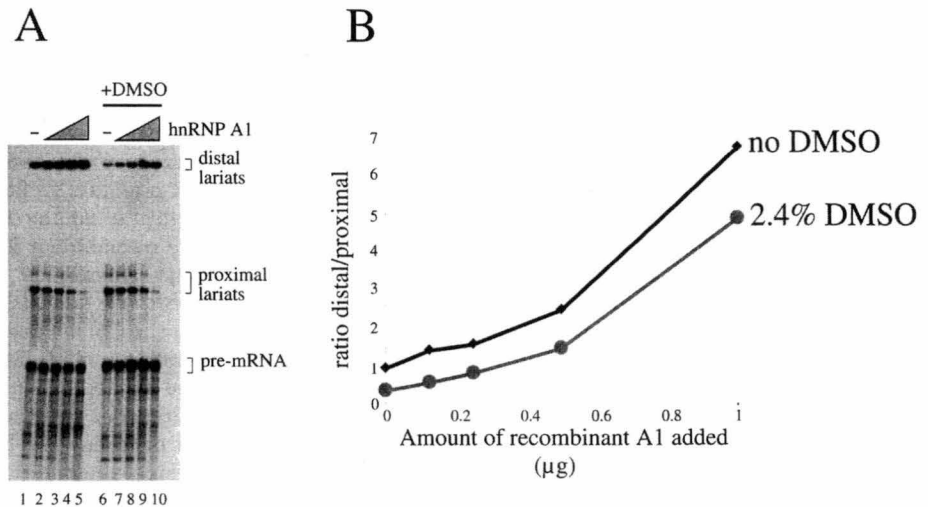
FIG. 2. Me_2SO affects 5'-splice site selection *in vitro*. **A**, structure of the pre-mRNAs used to assay modulation of 5'-splice site selection. C5' -/- and C5' 4/4 have been described previously (29). The C5' 4/4 pre-mRNA contains two CE4 elements, which are binding sites for hnRNP A1. **B**, incubation of the pre-mRNAs in HeLa extracts was for 2 h in the presence of different percentages of Me_2SO (0, 0.8, 1.6, 2.4%). Labeled RNA products were fractionated on a denaturing 11% polyacrylamide gel. The position and structure of the proximal and distal lariat products are shown.

tored the inclusion frequency of alternative exon 7B following the treatment of HeLa cells for 5 h with 5% Me_2SO (Fig. 1B). Although the effect was less dramatic than for the NCAM pre-mRNA, Me_2SO treatment significantly improved the inclusion of exon 7B.

Me_2SO Affects Splice Site Selection *in Vitro*—To determine whether Me_2SO can modulate splice site selection directly, we tested the effect of adding Me_2SO to splicing reactions incubated in nuclear extracts prepared from HeLa cells. We used model pre-mRNAs derived from the hnRNP A1 alternative splicing unit (29). C5' -/- contains two competing 5'-splice sites and a unique 3'-splice site (Fig. 2A). C5' -/- is spliced almost exclusively to the proximal 5'-splice site (Fig. 2B, lane 1). In contrast, the presence of A1 binding elements in C5' 4/4 promotes efficient splicing to the distal 5'-splice site (lane 5). The addition of Me_2SO at a final concentration of 0.8, 1.6, and 2.4% did not affect the splicing efficiency of C5' -/- RNA, and 5'-splice site selection remained exclusively proximal (Fig. 2B, lanes 2–4). In contrast, Me_2SO promoted a strong reduction in the use of distal 5'-splice site in C5' 4/4 pre-mRNA (lanes 6–8). The highest concentration of Me_2SO (lane 8) produced a 5-fold decrease in the use of the distal 5'-splice site. In some experiments, the reduction in distal 5'-splice site use was accompanied by an increase in the production of lariat products derived from the proximal 5'-splice site (e.g. see Fig. 5A, lane 2).

The effect of Me_2SO on 5'-splice site selection was as strong on a pre-mRNA that was synthesized in the absence of cap analogue (data not shown). Thus, the reduction in distal 5'-splice site usage was independent of the cap structure at the 5'-end of the pre-mRNA. Me_2SO also affected 5'-splice site

FIG. 3. Me₂SO does not affect the activity of recombinant hnRNP A1. A, HeLa extracts lacking or containing 2.4% Me₂SO were supplemented with recombinant hnRNP A1 proteins (0.125, 0.25, 0.5, and 1 μ g) and splicing of the C5' 4/4 pre-mRNA was monitored. The position of the distal and proximal lariat products are indicated. B, diagram depicting the A1-mediated stimulation of distal 5'-splice site usage in extracts lacking or containing Me₂SO. The almost identical slopes suggest that the activity of recombinant A1 is not affected by Me₂SO.



selection in a model pre-mRNA carrying two copies of the 5'-splice site of exon 7 (data not shown). Identical effects were seen with Me₂SO solutions obtained from different suppliers, and the deionization of Me₂SO did not change its activity on 5'-splice site selection (data not shown). Transient exposure of nuclear extracts to Me₂SO (*i.e.* incubation in the presence of Me₂SO followed by dialysis) did not affect 5'-splice site usage (data not shown). Thus, Me₂SO needs to be present in splicing mixtures to affect splice site selection.

Because Me₂SO had a strong effect on the alternative splicing of a pre-mRNA carrying A1 binding elements (C5' 4/4), we asked whether Me₂SO compromised the activity of the hnRNP A1 protein. We have shown previously that hnRNP A1 promotes distal 5'-splice site utilization on this pre-mRNA (29). In nuclear extracts containing Me₂SO, the addition of hnRNP A1 efficiently shifted selection toward the distal 5'-splice site (Fig. 3A, lanes 6–10). The effect was specific because the addition of similar amounts of GST or gene 32 protein had no effect (data not shown). Notably, the profile of stimulation obtained with increasing amounts of recombinant A1 was similar to the profile obtained in a nuclear extract lacking Me₂SO (Fig. 3A, lanes 1–5; compare the slopes in Fig. 3B). Because the activity of recombinant hnRNP A1 is not compromised by the presence of Me₂SO, it is unlikely that Me₂SO affects the activity of the endogenous A1 proteins.

To address whether Me₂SO has a similar activity on 3'-splice site selection we tested a pre-mRNA (C3' -/-; Fig. 4A), which is spliced predominantly to the distal 3'-splice site (Ref. 29; Fig. 4B, lane 1). C3' -/- splicing was sensitive to increasing amounts of Me₂SO (Fig. 4B, lanes 2–4). At the highest concentration of Me₂SO (lane 4), more than 50% of splicing occurred at the proximal 3'-splice site. We also tested a derivative of C3' -/- in which the central portion was substituted for the 3'-splice site region and a portion of NCAM alternative exon 18 (NCAM3' RNA). Although splicing of NCAM3' RNA was less sensitive to Me₂SO than C3' -/-, Me₂SO promoted a stronger reduction in the use of the distal 3'-splice site than of the proximal 3'-splice site (Fig. 4B, lanes 5–9). Alternative splicing of a β -globin pre-mRNA carrying duplicated 3'-splice sites was also affected by Me₂SO (data not shown).

Me₂SO Activates SR Proteins—The effect of Me₂SO on splice site selection is reminiscent of the activity of SR proteins, which tend to activate splicing of the proximal pair of splice sites (32, 33). Although Me₂SO did not stimulate overall splicing activity in nuclear extracts (Fig. 5A, lanes 1 and 2), we asked whether Me₂SO could mimic the generic splicing activity of SR proteins. This activity was initially defined by the capacity of SR proteins to activate splicing in a HeLa S100 extract.

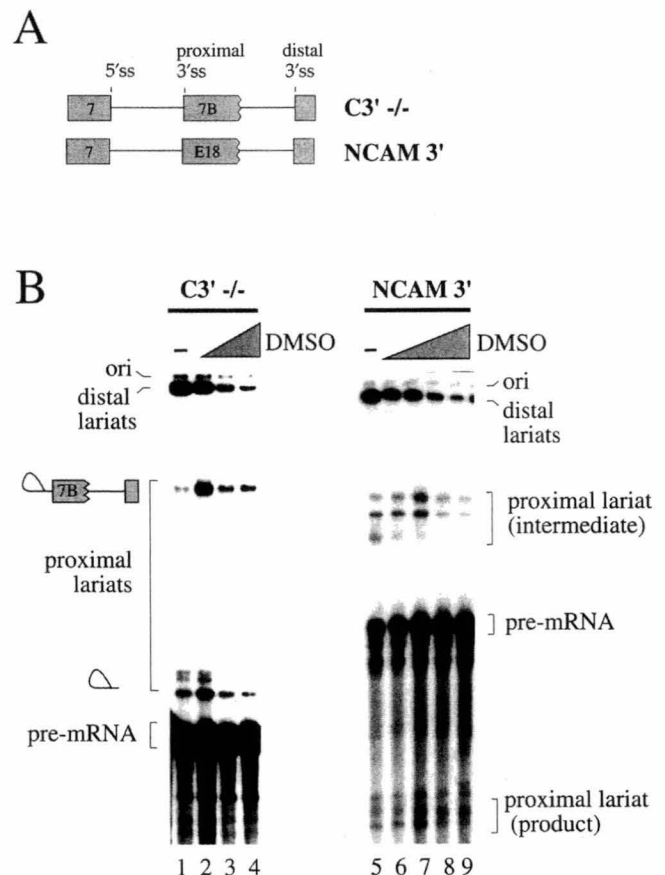


FIG. 4. Me₂SO affects 3'-splice site selection *in vitro*. A, structure of the pre-mRNAs used to assay modulation of 3'-splice site selection. C3' -/- is derived from the hnRNP A1 gene (29). The NCAM 3' pre-mRNA is a hybrid pre-mRNA containing the 5'-splice site of exon 7, the 3'-splice site of NCAM alternative exon 18 and the 3'-splice site of adenovirus L2 exon. B, *in vitro* splicing assays of model pre-mRNAs. Incubation was for 2 h in HeLa extracts containing increasing percentages of Me₂SO (0, 0.8, 1.6, 2.4% in lanes 1–4, and 0, 0.8, 1.6, 2.4, 3.2% in lanes 5–9). Labeled RNA products were fractionated on denaturing 11% (for C3' -/-) or 6.5% (for NCAM3') polyacrylamide gels. The position of the pre-mRNA and proximal and distal lariat products is shown.

either as a mixture of SR proteins or individually (33, 34). U2AF⁶⁵ also activates splicing when added to a HeLa S100 extract (35). Notably, the addition of Me₂SO to a HeLa S100 extract stimulated splicing as efficiently as the addition of the recombinant SR protein ASF/SF2 (Fig. 5A, lanes 3–5). The

addition of Me₂SO to a S100 extract also stimulated the formation of complexes, as judged by native gel analysis (Fig 5B, lanes 7–9). These results suggest that Me₂SO increases the activity of residual amounts of SR or U2AF proteins in the S100 extract. The level of splicing stimulation varied considerably in different preparations of S100 extract. Although Me₂SO and recombinant ASF/SF2 restored splicing activity in a similar manner, splicing to the distal 5'-splice site was not detected, as is the case in a nuclear extract (lanes 1 and 2). We have shown

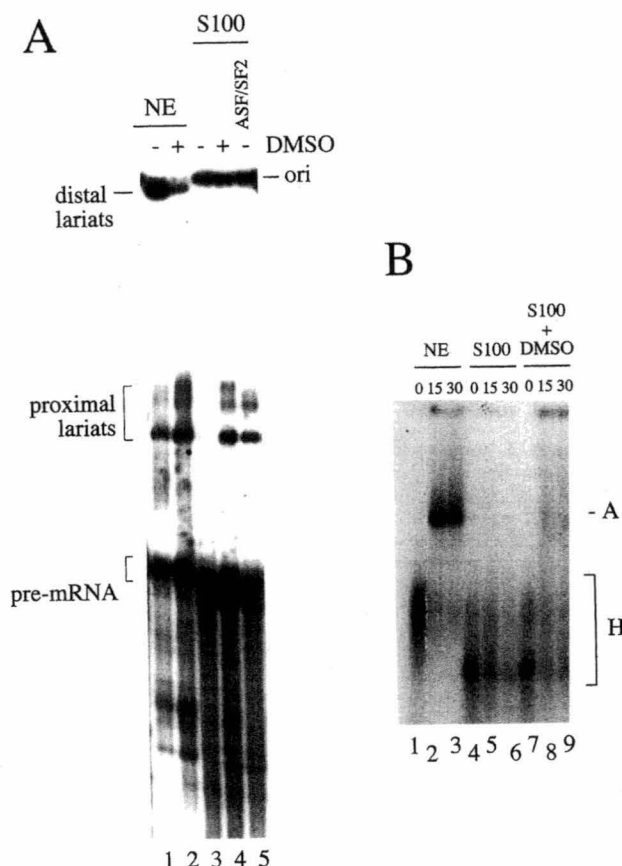


FIG. 5. Me₂SO rescues splicing in a HeLa S100 extract. A, splicing reactions were performed in a HeLa nuclear extract (NE) and in a HeLa S100 extract. The extracts were incubated in the absence or in the presence of 3.2% Me₂SO. The S100 extract was also supplemented with 0.5 μ g of recombinant ASF/SF2 protein (lane 5). The pre-mRNA substrate used was C5' 4/4. B, a HeLa nuclear extract (NE), a HeLa S100 extract or S100 extract supplement with 3.2% Me₂SO (S100 + DMSO) were incubated with a model pre-mRNA derived from the adenovirus major late transcription unit. Following incubation for the time indicated (in minutes), samples were fractionated in a native 5% polyacrylamide gel. The position of the nonspecific complex H and of splicing complex A is indicated.

previously that distal 5'-splice site selection on this pre-mRNA requires hnRNP A1 (29). The failure to activate distal 5'-splice site use is probably because of the fact that S100 extracts contain small amounts of hnRNP A/B proteins as compared with nuclear extracts (not shown).

The above result suggests that Me₂SO may affect the activity of SR proteins. To further examine this possibility, we tested the effect of adding Me₂SO to splicing reactions preincubated with a recombinant SR protein. At the concentrations used and in the absence of Me₂SO, the recombinant SR protein GST-SRp30c had little effect on 5'-splice site selection when using the C5' 4/4 pre-mRNA (Fig. 6A, lanes 1–3). However, in the presence of Me₂SO, the same amount of GST-SRp30c stimulated proximal 5'-splice site utilization (lanes 4–12). Thus, the simultaneous addition of Me₂SO and SR produced a shift toward proximal use that was greater than the sum of their individual contribution. Because recombinant SR proteins display more activity in the presence of Me₂SO, a similar effect on the endogenous SR proteins may be responsible for the activity of Me₂SO in nuclear extracts.

DMF and Formamide Also Modulate Splice Site Selection—To understand the chemical basis for the activity of Me₂SO in alternative splicing, we tested other solvents. At equivalent percentages, both DMF and formamide were at least as active as Me₂SO at modulating 5'-splice site selection (Fig. 7A). Surprisingly, although DMF and formamide shared with Me₂SO the ability to modulate 5'-splice site selection, formamide was unable to activate splicing in a HeLa S100 extract (Fig. 7B).

DISCUSSION

We have observed that the addition of Me₂SO to nuclear extracts can have strong effects on splice site selection while having minimal effects on the efficiency of splicing. In contrast, the addition of Me₂SO to a splicing-deficient HeLa S100 extract stimulated splicing in a manner reminiscent of the activity of SR proteins. The effect of Me₂SO on splice site selection was also similar to the activity of SR proteins because Me₂SO shifted selection toward the proximal pair of splice sites. Consistent with the notion Me₂SO stimulates the activity of SR proteins, we found that the combination of Me₂SO and SRp30c produces a shift that is greater than the sum of their individual contribution. Thus, a general stimulation in the activity of all endogenous SR proteins most probably explains why Me₂SO influences splice site choice *in vitro*. Likewise, the addition of Me₂SO to a S100 extract may stimulate the residual amounts of SR proteins present in this extract.

An additional mechanism by which Me₂SO may affect splice site selection is through the inactivation of the hnRNP A/B proteins, which are known to antagonize the activity of SR proteins in splice site selection (36, 37). However, we have

FIG. 6. Me₂SO activates SR proteins. A, using the C5' 4/4 pre-mRNA, the activity of the recombinant SR proteins GST-SRp30c was tested in the absence and in the presence of increasing concentrations of Me₂SO. The GST-SRp30c protein (0.5 and 1 μ g) was preincubated in nuclear extract 15 min at 30 °C before adding the pre-mRNA and Me₂SO. B, diagram depicting the SRp30c-mediated reduction in distal 5'-splice site usage in extracts lacking or containing different concentrations of Me₂SO.

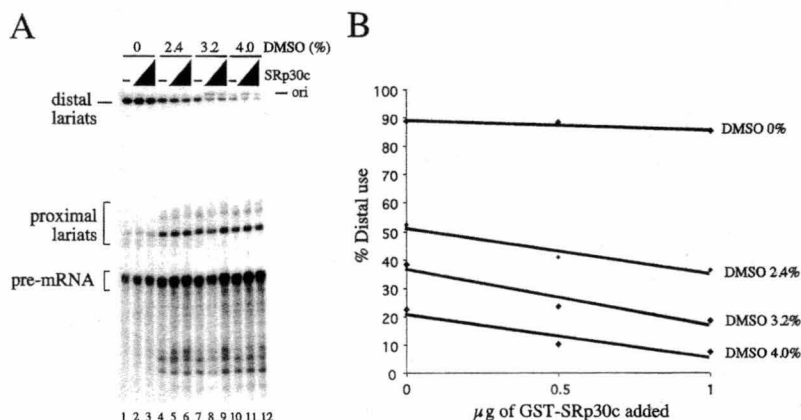
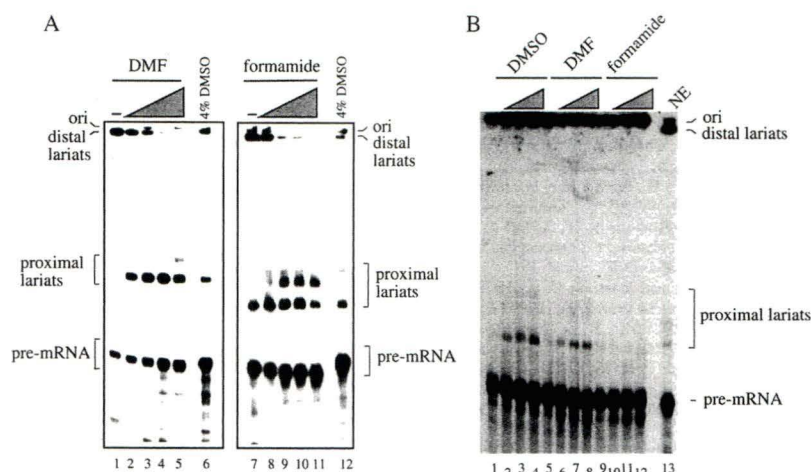


FIG. 7. Modulating and stimulating activities of DMF and formamide. A, *in vitro* splicing assays with the model pre-mRNA C5' 4/4 were carried out in HeLa extracts in the presence of different percentages of DMF or formamide (0, 1.6, 2.4, 3.2, and 4%). For comparison, splicing of the same pre-mRNA in a HeLa nuclear extract containing 4% Me₂SO is shown (lanes 6 and 12). The position of the pre-mRNA and proximal and distal lariat products is shown. B, splicing reactions with the C5' 4/4 pre-mRNA were set up in HeLa S100 extracts in the presence of increasing amounts of Me₂SO, DMF, or formamide (0, 2.4, 3.2, and 4%). For comparison, a splicing reaction performed in a HeLa nuclear extract is shown (NE, lane 13).



observed that the activity of recombinant hnRNP A1 proteins in splice site selection is not affected by the presence of Me₂SO. This result suggests that the activity of endogenous A1 proteins is probably not affected by Me₂SO.

Because Me₂SO, DMF, and formamide decrease the melting temperature of DNA and RNA duplexes, it is possible that changes in 5'-splice site selection may be caused by the weakening of the base pairing interactions between U1 snRNA and 5'-splice sites. However, this explanation is unlikely for the following reasons. First, a reduction in U1 snRNP binding should lead to more distal 5'-splice site utilization, as is the case when U1 activity is reduced through oligonucleotide-targeted RNase H digestion (data not shown). Moreover, increased selection of a proximal 5'-splice site is associated with improved U1 snRNP binding (38). Second, splice site selection on a transcript that is normally spliced to a distal 5'-splice site because of a duplex structure in the pre-mRNA was insensitive to Me₂SO.² Thus, the effect of Me₂SO, formamide and DMF on 5'-splice site selection cannot be explained by a reduction in the stability of base pairing interactions. However, it is possible that this denaturing activity contributes to the reduction in overall splicing activity observed at higher concentrations of solvents.

Although our results suggest that Me₂SO affects the activity of SR proteins, the mechanism by which SR proteins become activated remains unclear. Western analysis using an antibody that recognizes phosphorylated epitopes on SR proteins revealed no change in the overall and relative abundance of phosphorylated SR proteins upon incubation with Me₂SO (data not shown). Moreover, Me₂SO did not affect the binding of SR proteins to a purine-rich RNA splicing enhancer element (data not shown). Me₂SO also did not modify the solubility of SR proteins when extracts were incubated with increasing concentrations of MgCl₂ (data not shown). Although Me₂SO is regarded as a relatively inert solvent for pharmacological applications, it improves the solvation of cations and stimulates nucleophilic reactions. DMF and formamide share with Me₂SO this chemical property. Thus, one possibility is that Me₂SO improves the solvation of positive charges on proteins. This may influence the structure at the surface of proteins and facilitate ionic contacts between charged domains of interacting proteins. Consistent with this view, modulation of 5'-splice site selection *in vitro* is known to be sensitive to the ionic conditions of the reaction (39). Splicing proteins that carry charged domains include SR and U2AF proteins, which have RS-domains rich in positively and negatively charged amino

acids (arginines and phosphorylated serines, respectively). Interactions between the RS-domain containing proteins ASF/SF2, U1 snRNP-70 kDa, and U2AF³⁵ have been proposed to occur early during spliceosome assembly (40). Moreover, these interactions are sensitive to the phosphorylation state of ASF/SF2 (41). Thus, Me₂SO may activate splicing in a S100 extract by improving the quality of the ionic interactions between residual amounts of SR and U2AF proteins. Because the amount and activity of these proteins are in excess in a nuclear extract, this would explain why Me₂SO stimulates generic splicing in a S100 but not in a nuclear extract.

Even though nuclear extracts contain sufficient amounts of SR proteins for generic splicing, their activity in splice site selection is not maximal because adding more SR proteins can have a strong effect on the selection of splice sites (32, 33). Because a similar effect can be obtained by adding kinases that phosphorylate the RS domains of SR proteins (42), the profile of charged residues at the surface of SR proteins is crucial for their activity as splicing regulators. Moreover, the requirement for charged residues at the surface of SR proteins appears different for generic and alternative splicing because dephosphorylation of ASF/SF2 is essential for constitutive splicing, but is not required for the protein to function as a splicing regulator (41). Thus, Me₂SO may affect the presentation of charged residues that are important for the activity of SR proteins in splice site selection. Because Me₂SO and DMF activate splicing in a HeLa S100 extract, Me₂SO and DMF may also affect the presentation of different residues that are important for generic splicing. In contrast, because formamide affects splice site selection but cannot activate a S100 extract, formamide may only affect the presentation of residues that play a role in splice site selection.

Our results raise the possibility that the documented effect of Me₂SO on cell differentiation may be caused, at least in part, by changes in the activity of SR proteins, which in turn affect splice site selection. This conclusion is supported by the observation that DMF can mimic the effect of Me₂SO both in differentiation assays (5, 43–45), and in splicing assays *in vitro*. Depending on the cell types, Me₂SO can either promote or block differentiation or apoptosis. These opposite outcomes may be explained if different subsets of pre-mRNAs are expressed in different cell types. For example, alternative splicing is often used to control the production of proteins involved in programmed cell death such as Fas, Bcl-2, Bax, and Ced-4 (46). Hence, Me₂SO may alter the alternative splicing of a pre-mRNA to favor the production of a repressor protein in one cell type, while an inducer may be produced from another gene in a different cell type.

² M. Cordeau and B. Chabot, unpublished results.

Acknowledgments—We thank Jude Beaudoin for performing the RNase T1 protection assay on NCAM, and Johanne Toutant for help with cell culture. We thank Stephan Stamm for providing the GST-SRp30c plasmid, and Martin Simard for the SR protein preparations.

REFERENCES

- Yenofsky, R., Cereghini, S., Krowczynska, A., and Brawerman, G. (1983) *Mol. Cell. Biol.* **3**, 1197–1203
- Pollerberg, G. E., Schachner, M., and Davoust, J. (1986) *Nature* **324**, 462–465
- Nicholson, D. W., Ali, A., Klemba, M. W., Munday, N. A., Zamboni, R. J., and Ford-Hutchinson, A. W. (1992) *J. Biol. Chem.* **267**, 17849–17857
- Grunt, T. W., Somay, C., Oeller, H., Dittrich, E., and Dittrich, C. (1992) *J. Cell Sci.* **103**, 501–509
- Blau, H. M., and Epstein, C. J. (1979) *Cell* **17**, 95–108
- Fiore, M., and Degross, F. (1999) *Exp. Cell Res.* **251**, 102–110
- Wang, H., and Scott, R. E. (1993) *Cell Prolif.* **26**, 55–66
- Teraoka, H., Mikoshiba, M., Takase, K., Yamamoto, K., and Tsukada, K. (1996) *Exp. Cell Res.* **222**, 218–224
- Manduca, P., Castagnola, P., and Cancedda, R. (1988) *Dev. Biol.* **125**, 234–236
- Trubiani, O., Pieri, C., Rapino, M., and Di Primio, R. (1999) *Cell Prolif.* **32**, 119–129
- Marthyn, P., Beuscart, A., Coll, J., Moreau-Gachelin, F., and Righi, M. (1998) *Exp. Cell Res.* **243**, 94–100
- Melkonyan, H., Sorg, C., and Klempt, M. (1996) *Nucleic Acids Res.* **24**, 4356–4357
- Makowski, M., Ballester, R., Cayre, Y., and Rosen, O. M. (1988) *J. Biol. Chem.* **263**, 3402–3410
- Tam, S. P., Zhang, X., Cuthbert, C., Wang, Z., and Ellis, T. (1997) *J. Lipid Res.* **38**, 2090–2102
- Srinivas, S., Sironmani, T. A., and Shanmugam, G. (1991) *Exp. Cell Res.* **196**, 279–286
- Bahler, D. W., and Lord, E. M. (1985) *J. Immunol.* **134**, 2790–2798
- Campbell, P. L., Kulozik, A. E., Woodham, J. P., and Jones, R. W. (1990) *Genes Dev.* **4**, 1252–1266
- Pollerberg, E. G., Sadoul, R., Goridis, C., and Schachner, M. (1985) *J. Cell Biol.* **101**, 1921–1929
- Prentice, H. M., Moore, S. E., Dickson, J. G., Doherty, P., and Walsh, F. S. (1987) *EMBO J.* **6**, 1859–1863
- Pan, J. B., Monteggia, L. M., and Giordano, T. (1993) *Brain Res. Mol. Brain Res.* **18**, 259–266
- Emerit, M. B., Martres, M. P., Miquel, M. C., el Mestikawy, S., and Hamon, M. (1995) *J. Neurochem.* **65**, 1917–1925
- Bendori, R., Resnitzky, D., and Kimchi, A. (1987) *Virology* **161**, 607–611
- Klinken, S. P., Holmes, K. L., Morse, H. C. 3rd., and Thorgeirsson, S. S. (1988) *Exp. Cell Res.* **178**, 185–198
- Eick, D. (1990) *Nucleic Acids Res.* **18**, 1199–1205
- Darling, D., Tavassoli, M., Linskens, M. H., and Farzaneh, F. (1989) *Oncogene* **4**, 175–179
- Chabot, B. (1994) in *RNA Processing, A Practical Approach* (Higgins, S. J. and Hames, B. D., eds) Vol. 1, pp. 1–29, IRL Press, Oxford
- Melton, D. A., Krieg, P. A., Rebagliati, M. R., Maniatis, T., Zinn, K., and Green, M. R. (1984) *Nucleic Acids Res.* **12**, 7035–7056
- Chabot, B., Blanchette, M., Lapierre, I., and La Branche, H. (1997) *Mol. Cell. Biol.* **17**, 1776–1786
- Blanchette, M., and Chabot, B. (1999) *EMBO J.* **18**, 1939–1952
- Dignam, J. D., Lebovitz, R. M., and Roeder, R. G. (1983) *Nucleic Acids Res.* **11**, 1475–1489
- Tacke, R., and Goridis, C. (1991) *Genes Dev.* **5**, 1416–1429
- Ge, H., and Manley, J. L. (1990) *Cell* **62**, 25–34
- Krainer, A. R., Conway, G. C., and Kozak, D. (1990) *Cell* **62**, 35–42
- Zahler, A. M., Lane, W. S., Stolk, J. A., and Roth, M. B. (1992) *Genes Dev.* **6**, 837–847
- Zamore, P. D., and Green, M. R. (1991) *EMBO J.* **10**, 207–214
- Mayeda, A., and Krainer, A. R. (1992) *Cell* **68**, 365–375
- Mayeda, A., Munroe, S. H., Caceres, J. F., and Krainer, A. R. (1994) *EMBO J.* **13**, 5483–5495
- Eperon, I. C., Ireland, D. C., Smith, R. A., Mayeda, A., and Krainer, A. R. (1993) *EMBO J.* **12**, 3607–3617
- Schmitt, P., Gattoni, R., Keohavong, P., and Stevenin, J. (1987) *Cell* **50**, 31–39
- Wu, J. Y., and Maniatis, T. (1993) *Cell* **75**, 1061–1070
- Xiao, S. H., and Manley, J. L. (1998) *EMBO J.* **17**, 6359–6367
- Prasad, J., Colwill, K., Pawson, T., and Manley, J. L. (1999) *Mol. Cell. Biol.* **19**, 6991–7000
- Pise, C. A., Newburger, P. E., and Holland, C. A. (1992) *J. Gen. Virol.* **73**, 3257–3261
- Shen, Q., Chada, S., Whitney, C., and Newburger, P. E. (1994) *Blood* **84**, 3902–3908
- Hosseini, N. M., Brattain, D. E., McKnight, M. K., and Brattain, M. G. (1988) *Exp. Cell Res.* **175**, 125–135
- Jiang, Z. H., and Wu, J. Y. (1999) *Proc. Soc. Exp. Biol. Med.* **220**, 64–72

ANNEXE II

Gagne, J. P., Hunter, J., Labrecque, B., Chabot, B., and Poirier, G. G. (2003). A proteomic approach to the identification of heterogeneous nuclear ribonucleoproteins as a new family of poly(ADP-ribose)-binding proteins. *Biochem J* 371, 331-340.

J'ai contribué à l'article : "A proteomic approach to the identification of heterogeneous nuclear ribonucleoproteins as a new family of poly (ADP-ribose)-binding proteins", en produisant les protéines recombinantes hnRNPs utilisées pour l'expérience dont les résultats sont présentés à la figure 8 : "pADPr-binding pattern of conserved hnRNPs domain".

A proteomic approach to the identification of heterogeneous nuclear ribonucleoproteins as a new family of poly(ADP-ribose)-binding proteins

Jean-Philippe GAGNÉ*, Joanna M. HUNTER*, Benoît LABRECQUE†, Benoît CHABOT† and Guy G. POIRIER*¹

*Health and Environment Unit, Laval University Medical Research Center, Centre hospitalier universitaire de Québec (CHUQ), Faculty of Medicine, Laval University, 2705, Boulevard Laurier, Ste-Foy, Québec, Canada G1V 4G2, and †Department of Microbiology and Infectiology, Faculty of Medicine, Sherbrooke University, 3001, North 12th Avenue, Fleurimont, Québec, Canada J1H 5N4

A new class of poly(ADP-ribose) (pADPr)-binding proteins, heterogeneous nuclear ribonucleoproteins (hnRNPs), has been identified by a proteomic approach using matrix-assisted laser-desorption-ionization time-of-flight ('MALDI-TOF') MS. Liquid-phase isoelectric focusing with a Rotofor® cell (Bio-Rad) allowed pre-fractionation of proteins extracted from HeLa cells. Rotofor® protein fractions were further separated by SDS/PAGE and then transferred to a PVDF membrane. pADPr-binding proteins were analysed by autoradiography of the protein blot after incubation with ³²P-labelled automodified pADPr polymerase-1 (PARP-1). Peptide mass fingerprinting of selected bands identified the most abundant pADPr-binding proteins as hnRNPs, a family of proteins that bind pre-mRNA into functional complexes involved in mRNA maturation and transport to the cytoplasm. Sequence homology database searching against

a previously reported pADPr-binding sequence motif revealed that the hnRNPs contain a putative pADPr-binding sequence pattern [Pleschke, Kleczkowska, Strohm and Althaus (2000) J. Biol. Chem. 275, 40974–40980]. pADPr-binding assays performed with synthetic peptides by the dot-blot technique and with nitrocellulose-transferred recombinant hnRNPs confirmed the pADPr-binding protein identification and the specificity of the interaction. These results could establish a link between increased levels of pADPr in DNA damaged cells and the modified protein expression pattern resulting from altered mRNA trafficking.

Key words: heterogeneous nuclear ribonucleoprotein (hnRNP), peptide mass fingerprinting, poly(ADP-ribose) (pADPr), proteomic.

INTRODUCTION

Poly(ADP-ribosyl)ation is a unique post-synthetic modification catalysed by the growing poly(ADP-ribose) (pADPr) polymerase (PARP) family of enzymes in response to DNA-damaging agents [1,2]. Using NAD⁺ as a substrate, PARPs catalyse the formation of pADPr with heterogeneous branched chain lengths of up to 200 units *in vitro* [3]. Some of these enzymes, such as PARP-1 and PARP-2, are localized in the nucleus and some, such as PARP-4 and PARP-5, are located in the cytoplasm, but pADPr synthesis takes place mainly in the cell nucleus [3]. Recently, it has been shown that PARP-1 activation is required for translocation of apoptosis-inducing factor ('AIF') from the mitochondria to the nucleus and that PARP-1-dependent cell death is caspase-independent [4]. Also, post-translational poly(ADP-ribosyl)ation of an oscillator component has been proposed to contribute to setting the period length of the *Arabidopsis thaliana* (thale cress) central oscillator [5].

More than 30 nuclear substrates of PARPs have been identified [6]. However, the main acceptor of pADPr is PARP-1 itself, a phenomenon known as 'PARP-1 automodification' [7]. In view of the tremendous negative charge and size of pADPr, poly(ADP-ribosyl)ation of proteins likely has a major impact on protein structure and function.

In the present paper we report a proteomic approach to identify some proteins that bind to pADPr. The most abundant

of these pADPr-binding proteins are predominantly heterogeneous nuclear ribonucleoproteins (hnRNPs), a large set of proteins that bind mRNA in the cell nucleus. Although hnRNPs presumably have several different functions, these are not yet all well characterized for each member of the hnRNP family. It has been proposed that hnRNPs help package the pre-mRNA into functional complexes [8]. For example, several hnRNPs are associated with the spliceosome complex [9]. In addition to RNA binding, some hnRNPs are known to bind single-stranded, duplex and triplex DNA [10]. hnRNPs are multifunctional proteins that can influence gene expression through many cellular processes. For example, interaction between hnRNP A1 and inhibitory $\kappa B\alpha$ ('I $\kappa B\alpha$ ') is required for maximal activation of nuclear factor κB ('NF- κB ')-dependent transcription [11]. There is also considerable evidence that suggests a role for some hnRNPs in the export of mRNA from the nucleus to the cytoplasm [12–14]. The recent finding of telomere-bound hnRNPs suggests that hnRNPs may be involved in the regulation of telomere length [15].

The functions of conserved domains of hnRNP A1 and other hnRNPs have been extensively studied. hnRNP A1 is a 320-amino-acid protein containing two copies of an RNA recognition motif (RRM) at the N-terminus, known as the UP1 domain, and a C-terminal glycine-rich domain (CT) that comprises an M9 nuclear localization signal (NLS) involved in bidirectional nuclear–cytoplasmic transport [16] (Figure 1). RRM is common and

Abbreviations used: CCD, charge-coupled device; CT, C-terminal glycine-rich domain; DHBB, dihydroxyboryl-Bio-Rex; DTT, dithiothreitol; GST, glutathione S-transferase; HCCA, α -cyano-4-hydroxycinnamic acid; hnRNP, heterogeneous nuclear ribonucleoprotein; MALDI-TOF MS, matrix-assisted laser-desorption-ionization time-of-flight MS; NLS, nuclear localization signal; pADPr, poly(ADP-ribose); PARP, poly(ADP-ribose) polymerase; RO60, ribonucleoprotein 60; RRM, RNA recognition motif; TFA, trifluoroacetic acid.

¹ To whom correspondence should be addressed (e-mail guy.poirier@crchu.ulaval.ca).

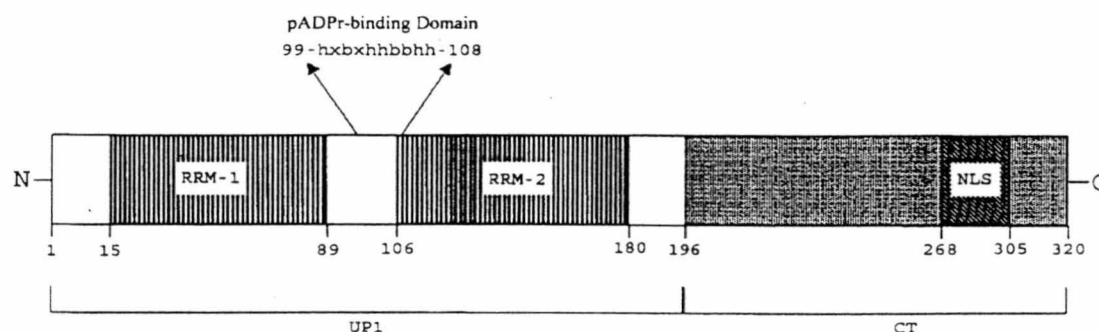


Figure 1 Structure of human hnRNP A1 conserved domains

Arrows indicate the putative consensus pADPr-binding domain. Numbers refer to the amino acid positions from the initiation codon. The positions of RRM and NLS are those specified by Vitali et al. [41] and Izaurralde et al. [16] respectively. The UP1 and CT domains are also indicated below the structure.

evolutionary conserved RNA-binding modules. The two RRM of hnRNP A1 are closely related, but have distinct functions in regulating pre-mRNA processing. The UP1 and C-terminal domains of hnRNP A1 can independently bind nucleic acids, although this binding is substantially weaker than that observed with intact hnRNP A1 protein [17]. Both the UP1 and the C-terminal domains of hnRNP A1 are required in alternative splicing of mRNA [18].

Recent reports have also suggested the implication of hnRNPs during apoptosis. The cleavage of hnRNP A1 by caspase-3 identified hnRNPs as apoptosis-associated proteins in human Burkitt lymphoma cell line [19]. Subcellular proteome analysis and differential protein expression of apoptotic cells has been very informative regarding hnRNP processing resulting from apoptotic events in the cell [19,20]. Little is known about the physiological significance of the involvement of hnRNPs in apoptosis, but translocation and cleavage of several other hnRNPs have been identified [21].

In the present study we have shown that hnRNPs which bind to pADPr contain a consensus sequence that is responsible for the specific binding and that pADPr binding is a characteristic distinct from RNA binding. The implications of these interactions in mRNA maturation and translocation will be discussed.

MATERIALS AND METHODS

HeLa cell sample preparation

The human cervical adenocarcinoma cell line HeLa-S3 was cultured in suspension (air/CO₂, 19:1; 37 °C) in Eagle's minimal essential medium (S-MEM; Sigma) containing 10% (v/v) fetal-bovine serum, 100 units/ml penicillin and 100 µg/ml streptomycin. Cell growth was produced in a sterile culture cell spinner. Cell pellets containing 2.6×10^8 cells were washed twice with sterile HEPES buffer and stored at -80 °C. They were then resuspended in 15 ml of the Rotofor® isoelectric-focusing (IEF) buffer [0.2% Triton X-100/5 M urea/2 M thiourea/2% CHAPS/2% Bio-Lytes pH 3-10 (Bio-Rad)/protease-inhibitor cocktail (according to Boehringer's instructions)]. The cell extract was vortex mixed for 5 min to achieve complete cell disruption and protein solubilization, then sonicated at 20 kHz with an amplitude value of 3 for 1 min with a 550 Sonic Dismembrator sonicator (Fisher Scientific). The cell extract was centrifuged at 15000 g for 15 min to remove extraneous cellular debris. The

volume of the cell-protein-extract supernatant was then adjusted to 18 ml with the Rotofor® buffer.

Liquid-phase IEF

A mini Rotofor® preparative cell (Bio-Rad) was used to separate proteins in the liquid phase according to their pI. The solubilized cell pellet was loaded into the Rotofor® focusing chamber without further treatment. A constant power of 10 W was applied to the system, which was cooled to 10 °C with a water circulator. The initial voltage of 450 V reached a plateau of 1800 V after approx. 4 h of focusing time. A total of 20 fractions were collected, and their pH values measured with a microelectrode. Each of the Rotofor® fractions were analysed on 1.5 mm-thick SDS/10%-(w/v)-polyacrylamide gels, which were silver-stained as described by Mortz et al. [22].

pADPr-binding protein identification

Even-numbered Rotofor® fractions were separated by SDS/10%-PAGE. Proteins were transferred on to a 0.45 µm-thick PVDF membrane (Bio-Rad) overnight at a constant voltage of 30 V in a system cooled to 4 °C. The membrane was incubated for 1 h at room temperature with gentle agitation in TBS-T [10 mM Tris/HCl (pH 8.0)/150 mM NaCl/0.1% Tween-20], containing 250 nM of either ³²P-labelled automodified PARP-1 or free ³²P-labelled pADPr. It was then washed with the TBS-T buffer until no radioactivity could be detected. The membrane was subsequently air-dried and subjected to autoradiography on Bio-Max (Kodak) films.

³²P-labelled automodified PARP-1 synthesis

³²P-labelled automodified PARP-1 was synthesized essentially as described by Ménard and Poirier [23]. In a total reaction volume of 900 µl [100 mM Tris/HCl (pH 8.0)/10 mM MgCl₂/8 mM dithiothreitol (DTT)/10% (v/v) glycerol/calf thymus activated DNA (23 µg)/1 mM NAD⁺], 75 µCi of [³²P]NAD⁺ was added. Ethanol was added to this preparation dropwise at 10% (v/v) final concentration, with constant mixing, and the reaction mixture was incubated for 3 min at 30 °C. The reaction was started by adding 20 units [24] of PARP-1 purified up to the

DNA-cellulose step (600 units/mg of protein) as described by Zahradka and Ebisuzaki [24]. After 30 min at 30 °C, during which time the enzyme was modified by covalent linkage of pADPr chains to its automodification domain, 100 μ l of 3 M sodium acetate, pH 5.2, and 700 μ l of propan-2-ol were added as described by Brochu et al. [25]. The reaction mixture was kept on ice for 30 min and then centrifuged at 10 000 *g* for 10 min at 4 °C. The pellet was washed twice with ice-cold 80 % (v/v) ethanol and resuspended in 900 μ l of TE buffer [10 mM Tris/HCl (pH 8.0)/1 mM EDTA]. Calculating from the radioactivity count before and after synthesis, the percentage of NAD⁺ incorporation was 16 %. The final pADPr concentration was 162 μ M.

³²P-labelled free pADPr

Following ³²P-labelled automodified PARP-1 synthesis, the protein-bound pADPr was hydrolysed by incubation with 1 ml of 1 M KOH and 50 mM EDTA solution at 60 °C. After 1 h, 9 ml of AAGE9 buffer [250 mM ammonium acetate (pH 9.0)/6 M guanidine/10 mM EDTA] was added and the pH was adjusted to 9.0 with 4 M HCl. The free pADPr was isolated on a dihydroxyboryl Bio-Rex (DHBB) resin synthesized as described by Shah et al. [26]. A 0.5 ml portion of DHBB was packed into a 10 ml Econo column (Bio-Rad) and washed with 5 ml of water and 10 ml of AAGE9 buffer. The sample was loaded on the DHBB, the flow-through was reloaded, and the resin was washed with 20 ml of AAGE9 buffer, followed by 10 ml of 1 M ammonium acetate, pH 9.0. The free ³²P-labelled pADPr was eluted with five portions of 0.5 ml of water at 37 °C. Fractions 2 and 3, corresponding to the peak of radioactivity, were pooled.

Identification of pADPr-binding proteins by peptide mass fingerprinting

Prior to the protein extraction and identification on the matrix-assisted laser-desorption-ionization time-of-flight (MALDI-TOF) mass spectrometer, the ³²P-labelled pADPr was removed from the proteins by incubating the PVDF membrane for 30 min in a 0.1 M NaOH solution containing 20 mM EDTA. Bands excised from the PVDF membrane were rinsed with de-ionized water and cut into pieces of approximately 1 × 1 mm². To each band, 100 μ l of 25 mM ammonium bicarbonate in 10 % acetonitrile was added. Proteins in the PVDF membrane pieces were reduced by adding 3 μ l of 45 mM DTT and incubating for 20 min at 37 °C. Alkylation was accomplished by adding 3 μ l of 100 mM iodoacetamide and incubating the mixture in the dark for 20 min. This buffer was then removed, and 100 μ l of 25 mM ammonium bicarbonate in 10 % acetonitrile and 0.1 % octyl β -D-glycopyranoside was added. The samples were incubated at 37 °C for 20–24 h with 5 μ l of 0.1 mg/ml trypsin (Promega Sequencing Grade Modified Trypsin).

An α -cyano-4-hydroxycinnamic acid (HCCA) matrix solution was prepared as a saturated solution in 50 % acetonitrile and 0.1 % trifluoroacetic acid (TFA). A 1 μ l aliquot of the digestion solution was pipetted directly on to a dried 0.8 μ l spot of HCCA on the sample plate. The spot was allowed to dry and then another 1 μ l was deposited. Finally, the crystallized sample spot was washed twice with 2 μ l of 0.1 % TFA. An Applied Biosystems Voyager-DETM PRO BiospectrometryTM Workstation was used to acquire positive-ion mass spectra. Data were collected in reflectron mode at an acceleration voltage of 20 kV and an extraction delay time of 200 ns. Spectra were internally calibrated using the [*M* (the molecular ion) + H]⁺ peaks at 842.502 and 2211.097 Da resulting from trypsin autoproteolysis.

Peptide masses measured by MALDI-TOF MS were submitted to PeptIdent (ExPASy Molecular Biology server) for searching against the SwissProt database or to ProFound (Proteometrics) for searching the NCBI (National Center for Biotechnology Information, U.S. National Library of Medicine Bethesda, MD, U.S.A.) non-redundant protein database. The mass-tolerance parameter for matching monoisotopic [*M* + H]⁺ ion masses was set at 30 p.p.m. Protein molecular masses were constrained to ± 100 % of the measured molecular mass, and pI values were not restricted (0–14 pH units). A maximum of 1 missed trypsin cleavage was allowed in the search. Results were evaluated and prioritized based on sequence coverage and matching of the theoretical to the measured molecular mass and pI.

Dot-blot analysis of the pADPr-binding domain

The amino acid sequences of histone H4, ribonucleoprotein 60 (RO60), hnRNP A1 and hnRNP A2 were searched for sequence homology with the pADPr-binding sequence motif given by Pleschke et al. [27]. Searching for this binding pattern was accomplished using the ScanProsite tool on the ExPASy Molecular Biology server. Peptides corresponding to the putative pADPr-binding domain of these proteins as well as alanine-substituted peptide sequences from the original hnRNP A2 were synthesized using fluorenyl-methoxycarbonyl ('Fmoc') technology on an Applied Biosystems 433A peptide synthesizer. For the dot-blot analysis, the following peptides were used:

histone H4: KRHRKVLDRDNIQGITKPAIRRLARR

RO60: FKKDLKESMKCGMWGRALRKA

hnRNP A1: REDSQRPGAHLTVKKIFVGGIK

hnRNP A2: REESGKPGAHVTVKKLFVGGIK

hnRNP A2 substitution #1:

AAAAAAPGAHVTVKKLFVGGIK

hnRNP A2 substitution #2:

REESGKPGAAVTVAAALFVGGIA

hnRNP A2 substitution #3:

REESGKPAAHATAKKAFFVGGIK

hnRNP A2 substitution #4:

REESGKPGAHVTVKKAAAAAAK

negative control 1: CHRPLMRNQKSRDSS

negative control 2: HEGVYIEPEARGLRC

The dot-blot was performed essentially as described by Panzeter et al. [28]. Peptides were dissolved in TBS-T and 2 μ g of each peptide was spotted on to a 0.05 μ m nitrocellulose membrane (Schleicher und Schüll). The membrane was air-dried, rinsed three times with TBS-T and incubated for 1 h at room temperature (22 °C) with gentle agitation in TBS-T containing 250 nM ³²P-labelled automodified PARP-1 or free ³²P-labelled pADPr. It was then washed with TBS-T buffer until no radioactivity could be detected. The membrane was subsequently air-dried and subjected to autoradiography on Bio-Max (Kodak) films.

Immunological detection of pADPr-binding peptides

The TBS-T washing of the membrane was followed by a 1 h incubation at room temperature with TBS-T containing 250 nM free non-radioactive pADPr. The membrane was rinsed three

times with TBS-T and incubated in PBS-MT (PBS buffer containing 5% non-fat dried milk and 0.1% Tween-20) for 1 h. The anti-pADPr antiserum 96-10 (1:10000 in PBS-MT) [29] was applied overnight with gentle agitation at room temperature. The membrane was washed several times with PBS-MT and incubated for 30 min with a secondary anti-rabbit antibody (1:2500 in PBS-MT) conjugated with peroxidase (Jackson ImmunoResearch Laboratories). The membrane was washed and revealed using the Renaissance chemiluminescence kit (NEN Life Sciences, now PerkinElmer Life Sciences). One dot-blot control membrane was stained with SYPRO™ Ruby Red stain (Molecular Probes) in order to evaluate the uniformity of peptide retention on the membrane, as displayed by the integrated density value of the fluorescence intensity on a UV-based charge-coupled-device (CCD) camera (Alpha Innotech Corporation, San Leandro, CA, U.S.A.).

SYPRO™ Ruby Red protein blot stain

The nitrocellulose membrane was fixed in 10% (v/v) methanol and 7% (v/v) acetic acid for 15 min. The membrane was rinsed three times with water and covered for 30 min, with gentle agitation, with SYPRO™ Ruby Protein blot stain. The blot was rinsed for 15 min with water and revealed using a fluorescence-based CCD imaging system (Alpha Innotech Corporation).

Expression and purification of hnRNP conserved domains

Plasmids pGEX-A1 and pGEX-A2 were constructed by inserting the mouse hnRNP A1 and hnRNP A2 cDNAs [30] into the *EcoRI* site of pGEX-2T. The hnRNP UP1 fragment was produced as described in LaBranche et al. [31] and was inserted into the *EcoRI* site of pGEX-2T. pGEX-hnRNP A1ΔRRM2 was produced by the deletion of the *BglII*–*MscI* fragment and contains a linker element. hnRNP A1ΔRRM2 therefore lacks amino acids 107–171, but contains instead amino acids Ile-Ser-Ser. The fragment hnRNP A1-CT containing amino acids 194–319 was produced by PCR and cloned in pGEX-2T.

All recombinant glutathione S-transferase (GST)–hnRNP fusion proteins were expressed in *Escherichia coli* BL21. Bacteria were grown for 2–3 h at 37 °C until reaching an attenuation (D_{600}) between 0.5 and 1.0. Induction with 100 μ M isopropyl 1-thio- β -D-galactopyranoside ('IPTG') was performed for 4 h. Cells were washed with cold Rec-Buffer [50 mM piperazine/HCl (pH 9.8)/500 mM NaCl/1 mM DTT/1 mM EDTA/0.5 mM PMSF/1 mM benzamidine/bacitracin (20 mg/ml)] and resuspended with 10 ml of Rec-Buffer/100 ml of culture. Lysozyme (0.4 mg/ml) was added 30 min before sonication. Cells were sonicated for 15 s and left on ice for 30 s. The cycle was repeated another three times. Triton X-100 was added to a final concentration of 1%, and the mixture was incubated for 30 min at room temperature on a rotating device. After a 10 min centrifugation at 12000 *g* at 4 °C, 600 ml of GSH–Sephadex (from a 50% slurry in Rec-Buffer) was added to the supernatant. The mixture was incubated for 30 min at 4 °C, then centrifuged 20 s at 500 *g* and washed four times with Rec-Buffer containing 0.1% Triton X-100. GST-proteins were eluted with 300 ml of a buffer comprising 20 mM GSH, 200 mM piperazine, pH 9.8, 1 mM DTT, 1 mM EDTA and 500 mM NaCl. Eluted material was dialysed in Buffer D [20 mM Hepes/KOH (pH 7.9)/100 mM KCl/20% (v/v) glycerol/0.2 mM EDTA/1 mM DTT/0.5 mM PMSF]. A 1 μ g portion each of recombinant hnRNPs, core histones and GST-protein was separated by electrophoresis on 0.75-mm-thick SDS/12% polyacrylamide gel. The proteins were transferred on to a 0.05- μ m-thick nitrocellulose membrane (Schleicher and

Schüll). A constant voltage (125 V) was applied for 1 h to the system, which was cooled to 4 °C. Membranes were washed three times with TBS-T before incubation with 32 P-labelled auto-modified PARP-1. Competition assays using RNA and NAD $^{+}$ were also performed in order to evaluate specificity of pADPr binding to hnRNPs. A 100-fold excess of yeast RNA (Ambion) or NAD $^{+}$ (Boehringer Mannheim) over pADPr was added to the incubation buffer. Again, the membranes were washed until no radioactivity could be detected in the TBS-T buffer. On occasions, membranes were washed under high-stringency conditions (1 M NaCl in TBS-T). The membranes were air-dried and subjected to autoradiography on Bio-Max (Kodak) films.

RESULTS

Identification of pADPr-binding proteins

Liquid-phase IEF followed by SDS/PAGE was used to fractionate proteins from a HeLa-cell protein extract. Pre-fractionation by liquid-phase IEF decreases considerably the sample complexity prior to SDS/PAGE. Moreover, preparative IEF permits much higher protein load than does a conventional two-dimensional gel, ranging up to the milligrams of protein. Twenty Rotofor® fractions of HeLa-cell proteins were collected over an almost linear pH range from 2 to 10 (Figure 2). Rotofor® fractions analysed by SDS/PAGE (Figure 3) illustrate that most of the proteins do not focus in only one pH fraction, but rather spread over two or more pH fractions with a clear enrichment pattern in some of these fractions. The goal of this two-step separation is to minimize the number of co-migrating proteins. Although MALDI-TOF MS can deal with multiple protein identification within a band, it is desirable to obtain a single protein per band.

After transfer of the PAGE-separated proteins to a PVDF membrane and exposure to 32 P-labelled automodified PARP-1, autoradiography reveals the putative pADPr-binding proteins (Figure 4). On comparison of the very different patterns obtained from silver-stained gel and autoradiography obtained with 32 P-labelled automodified PARP-1, one can observe that cellular proteins do not bind the pADPr in proportion to their abundance. As expected, the pADPr-binding patterns obtained from SDS/

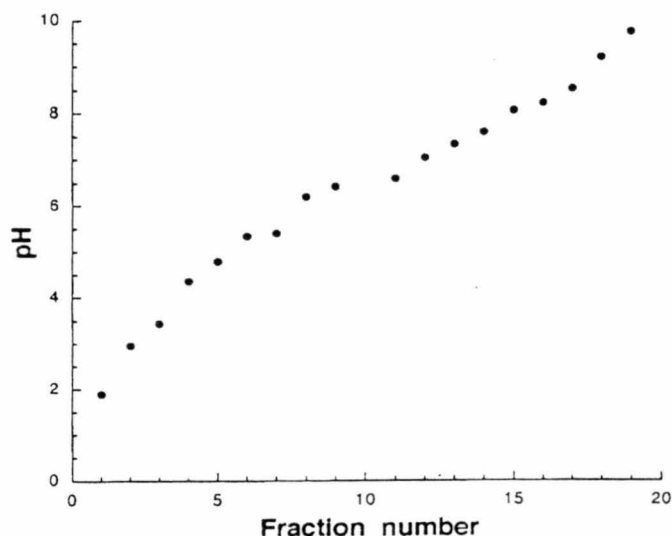


Figure 2 pH evaluation of Rotofor® fractions harvested from a pH 3–10 electrofocusing chamber

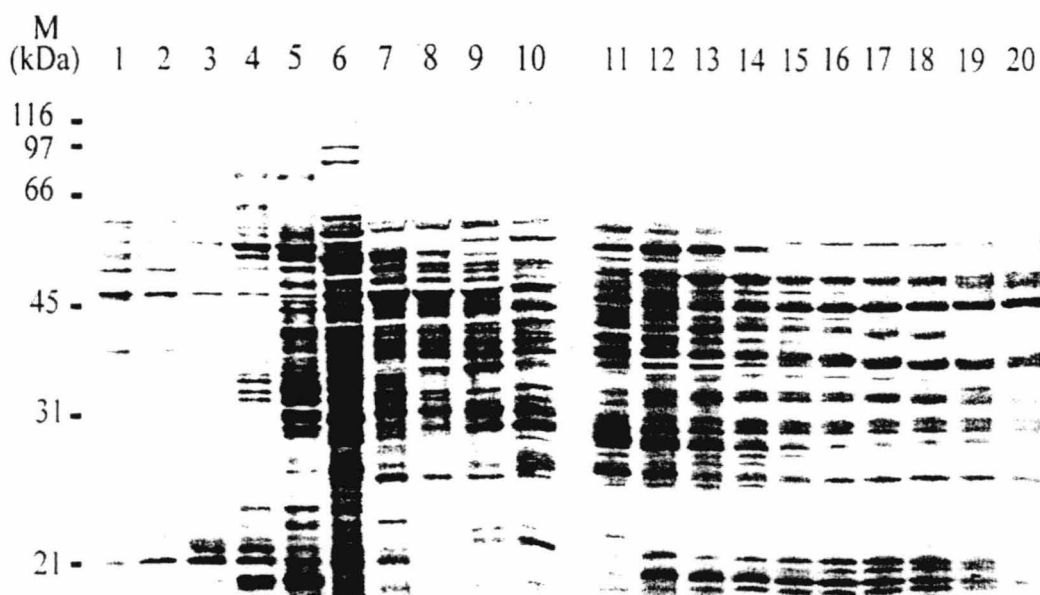


Figure 3 Silver-stained SDS/PAGE of Rotofor® fractions

The fraction numbers are indicated over each lane. M is molecular mass.

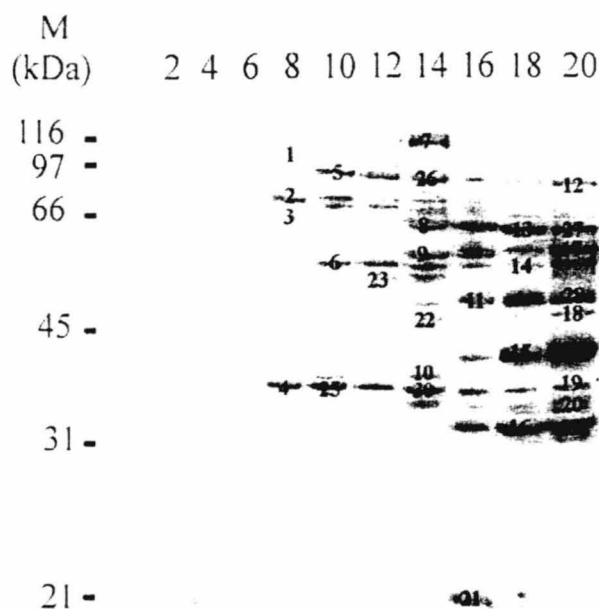


Figure 4 Analysis of pADPr-binding proteins

Even-numbered fractions harvested from the Rotofor® cell were transferred on to PVDF membrane. Incubation with 32 P-automodified PARP-1 or free [32 P]pADPr followed by autoradiography reveals a binding-protein pattern. Numbered bands were excised for identification by peptide mass fingerprinting. Rotofor® fraction numbers are indicated over each lane. M is molecular mass.

PAGE-separated Rotofor® fractions using either automodified PARP-1 or free pADPr are identical (results not shown). This observation indicates that the interaction is a protein-pADPr

contact rather than a protein-protein interaction. Considering that most of the cellular pADPr is protein-bound and that PARP-1 is the most important pADPr acceptor, the use of automodified PARP-1 is more likely to be representative of the interaction.

A total of 30 of the higher-intensity bands were cut from the membrane and analysed by peptide mass fingerprinting after trypsin digestion. The mass spectrum shown in Figure 5 is a tryptic digest from the band number 29 (see Figure 4) in which hnRNP A2/B1 and annexin II co-migrated. In this particular band, the use of recombinant hnRNP A2 showed that hnRNP A2 is actually the pADPr-binding protein, even if the relative intensity of hnRNP A2 peptides is considerably lower compared with annexin II. There were 15 tryptic peptides identified from hnRNP A2/B1 and eight peptides from annexin II. MALDI-TOF MS identification reveals many hnRNPs as pADPr-binding proteins (hnRNP A2/B1, hnRNP C1/C2, hnRNP E1, hnRNP G, hnRNP H, hnRNP K and hnRNP M) as well as previously reported pADPr-binding proteins such as lamins and histones (Table 1). Note that the majority of these proteins have a basic pI. The identity of the proteins in other bands will be the subject of further study. In general, no identifications were obtained from bands having a molecular mass higher than 90 kDa, probably because the higher-molecular-mass proteins are not efficiently transferred to the membrane. Most of the pADPr-binding proteins were localized in a single protein band, but some hnRNPs co-migrated with other proteins.

Characterization of pADPr binding on synthetic peptides by dot-blot analysis

In order to characterize further this pADPr affinity, we attempted to localize the pADPr-binding site of hnRNPs on the basis of a previously reported pADPr-binding consensus sequence [27] (see Figure 6). This consensus pattern consists of approx. 20 amino acids: a N-terminal cluster rich in basic residues (lysine/arginine) and a C-terminal region containing alternating hydrophobic and basic amino acids (Figure 6A). A search with the PATTINPROT

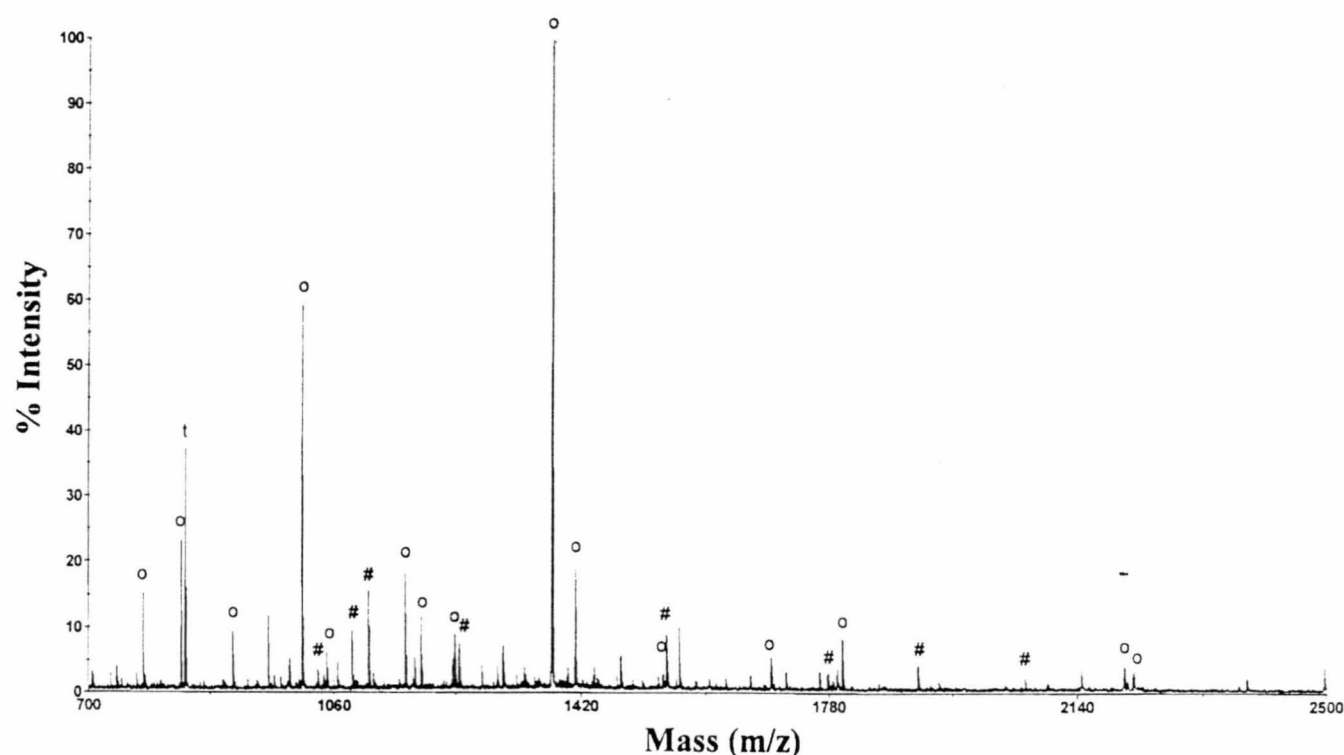


Figure 5 Tryptic-peptide mass map of the proteins in band number 29

Peptides from two proteins, hnRNP A2/B1 (#) and annexin II (°), were identified in this band. Trypsin autolysis products (t) are also labelled.

Table 1 pADPr-binding proteins from a HeLa-cell extract identified by peptide mass fingerprinting

Bands were excised from the membrane shown in Figure 4.

Band no.	Tentative protein identification	SwissProt or GeneBank* accession no.	Sequence coverage (%)	pI/molecular mass (Da)	
				Theoretical	Estimated
4	hnRNP C1/C2 (2 isoforms)	P07910	20	5.11/33 299; 5.10/31 966	6/39 000
6	hnRNP K (2 isoforms)	Q07244	19	5.39/50 976; 5.19/51 028	6.5/57 000
8	Lamin A/C	P02545	40	6.57/74 139	8/66 000
9	Lamin A/C (second isoform)	P02545	30	6.4/65 135	8/60 000
16	hnRNP A2/B1 (isoform) and Annexin II	P22626; P07355	37; 31	8.67/36 006; 7.56/38 473	9/35 000
20	hnRNP A2/B1	P22626	19	8.97/37 430	10/38 000
21	Histone H1C	P16402	22	11.02/22 219	8.5/20 000
22	hnRNP H	P31943	17	5.9/49 229	8/50 000
24	hnRNP G (RNA-binding motif protein) and phosphoglycerate kinase 1	P38159 (11421425); P00558 (6456828)	40; 37	10.02/42 404; 8.3/44 597	10/42 000
25	hnRNP C1/C2 (2 isoforms)	P07910	19	5.11/33 299; 5.10/31 966	6.5/39 000
27	hnRNP M (2 isoforms)	P52272	31	8.94/77 470; 9.03/73 575	10/66 000
29	hnRNP A2/B1 (2 isoforms) and annexin II	P22626; P07355	46; 25	8.97/37 430 or 8.67/36 006; 7.56/38 473	9/35 000
30	hnRNP C1/C2 (2 isoforms) and hnRNP E1 (α-CP1)	P07910; Q15365 (2134737)	25; 20	5.11/33 299 (5.10/31 966); 6.66/37 526	8/39 000

tool on the NPS@ server (Network Protein Sequence @analysis, <http://npsa-pbil.ibcp.fr/>) [32] revealed that all of the identified hnRNPs may contain such a pADPr-binding site, it being borne in mind that the pADPr-binding domain does not contain invariant amino acids at a given position within the consensus. Multiple sequence alignment of related hnRNPs with the consensus sequence reveals a closely matching pattern (Figure 6B).

To verify whether the putative pADPr-binding domain of hnRNPs does indeed bind to pADPr, a pADPr-binding dot-blot

analysis was performed on synthetic peptides that display the consensus sequence. The Western blot shown in Figure 7(B) clearly demonstrates that the putative pADPr-binding domain of histone H4, RO60, hnRNP A1 and hnRNP A2 mediates pADPr binding. The use of different approaches to characterize the pADPr-binding affinity of dot-blotted synthetic peptides results in approximately the same pattern of binding affinities. Using either radiolabelled automodified PARP-1 and free pADPr or the immunological detection of pADPr with the 96-10 antibody

A

hnRNP A0	REDSARPGAHAKVKKLFVGGIK
hnRNP A1	REDSQRPGAHLTVKKIFVGGIK
hnRNP A2	REESGKPGAHTVKKLFVGGIK
Consensus	bxxxxbxhxbx hbbhhhhhhb
pADPr-binding consensus	[xxxxK/Rxxx]xhxbxhbbhhbxxxx
hnRNP A2 Substitution #1	AAAAAAGAHTVKKLFVGGIK
hnRNP A2 Substitution #2	REESGKPGAATVVAALFVGGIA
hnRNP A2 Substitution #3	REESGKPAAHATAKKAFFVGGIK
hnRNP A2 Substitution #4	REESGKPGAHTVKKAAAAAAK

B

Identified hnRNPs	Pattern site	Similarity (%)
hnRNP A2/B1	106-GAHTVKKLF-115	93
hnRNP C1/C2	92-RGKAGVKRSA-101	86
hnRNP E1	25-VGSIIGKKGE-34	71
hnRNP G	118-GTRGPPSRGG-127	78
hnRNP H	210-YDRPGAGRGY-219	79
hnRNP K	261-FDRMPGRGG-270	78
hnRNP M	212-DYKVGWKKLK-221	86

Figure 6 (A) Sequence alignment of the putative pADPr-binding sequence for closely related hnRNPs corresponding to the dot-blot peptide synthesis and (B) putative pADPr-binding site and similarity with the consensus sequence for all hnRNPs identified by peptide mass fingerprinting

(A) The basic amino acids (b) are lysine, arginine and histidine, and the hydrophobic amino acids (h) are alanine, cysteine, glycine, valine, isoleucine, leucine, methionine, phenylalanine, tyrosine and tryptophan. Positions indicated by 'x' can be any amino acid. Positions in **bold** are consensus amino acids common to both sequences. The sequence of the hnRNP A2 peptides modified by replacement with alanine is aligned with the pADPr-binding consensus sequence. Substituted amino acid positions are underlined. (B) The given similarity percentages are those calculated from the PATTINPROT algorithm [32].

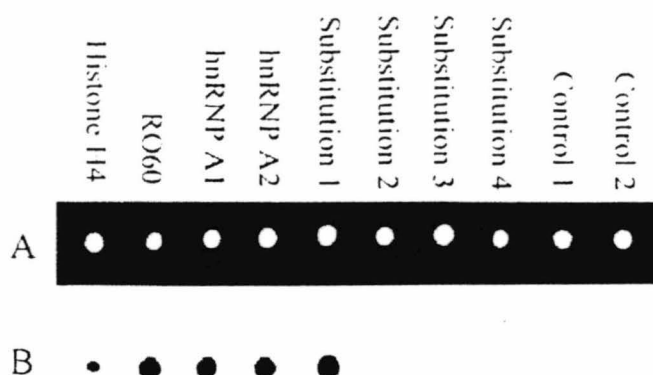


Figure 7 pADPr-binding dot-blot analysis

Different peptides that comprise the putative pADPr-binding site (see the text) were blotted on to nitrocellulose membrane. (A) SYPRO Ruby Red blot stain. (B) Immunological detection of pADPr-binding peptides in the presence of 250 nM automodified PARP-1. Detection was accomplished by using the anti-pADPr rabbit polyclonal antibody 96-10 in a Western-blot assay.

did not modify the observed pADPr-binding affinity. Moreover, the use of high-stringency buffers (TBS-T with 1 M NaCl) did not abolish pADPr-binding (results not shown). The immunological detection of protein-bound pADPr is presented in Figure 7.

In addition, we have verified by SYPRO Ruby Red blot staining that the quantity of each peptide retained on the

membrane was the same (Figure 7A). Uniformity of the peptide retention on the nitrocellulose membrane allows a direct comparison between the pADPr-binding efficiencies of each peptide. To obtain the same amount of each peptide bound to the nitrocellulose membrane, we evaluated the fluorescence intensity of a standardized blot stained with Sypro Ruby Red. The integrated density value of the fluorescence intensity for each of these peptides was approximately the same, indicating an equal amount of peptide retained on the membrane.

It has been proposed that the hydrophobic amino acids in the binding motif are necessary for pADPr binding, whereas the basic residues are less critical [27]. To test this hypothesis, the hnRNP A2 peptide sequence was selectively modified by alanine substitution of all basic and hydrophobic amino acids. The substituted peptides are listed in Figure 6(A). As shown in Figure 7, when the N-terminal tail of the consensus sequence is replaced by alanine (substitution 1), the peptide does not significantly decrease the binding to pADPr. Consequently the consensus could be localized to the hydrophobic/basic region of the consensus without the lysine/arginine-rich cluster at the N-terminus. Replacement by alanine of all the basic residues in the consensus sequence (substitution 2) completely inhibits pADPr binding, as does replacing the hydrophobic residues (substitution 3). Finally, replacement by alanine of the C-terminal residues also abolishes pADPr binding. That the two most hydrophobic residues of the putative pADPr-binding sequence (leucine and phenylalanine) are replaced with residues of much lower hydrophobicity could explain the complete loss of pADPr binding. In contrast with what has been proposed, both interspersed hydrophobic and basic amino acids seem to be important in achieving pADPr binding. Also, preliminary results with random-generated

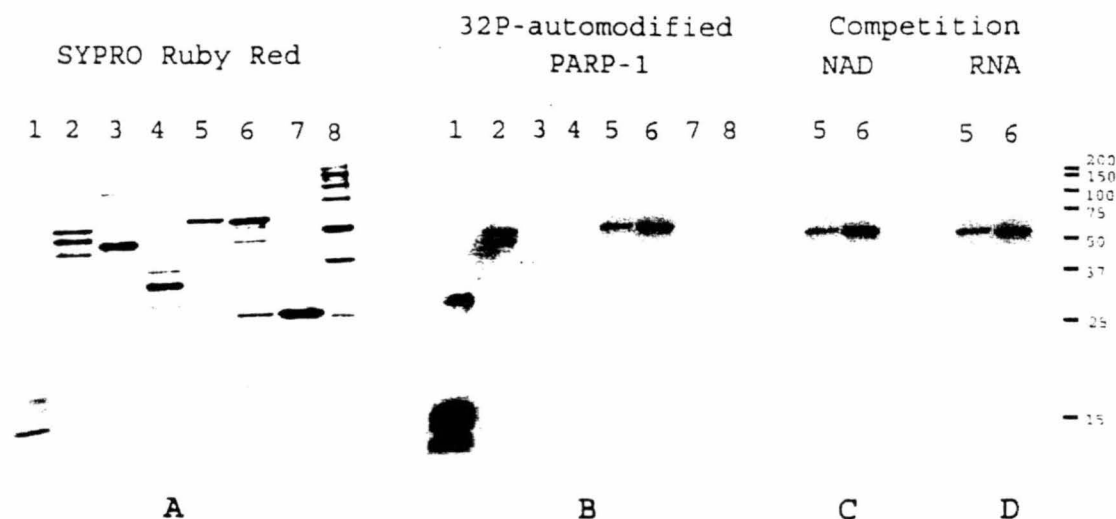


Figure 8 pADPr-binding pattern of the conserved hnRNP domain

Recombinant purified hnRNP protein constructs (1 μ g) were separated on SDS/12% PAGE, followed by transblot on to nitrocellulose membrane. Membranes were stained with SYPRO Ruby Red protein blot stain (A), incubated with 250 nM 32 P-automodified PARP-1 in TBS-T (B), 250 nM 32 P-automodified PARP-1 with a 100-fold excess of NAD $^{+}$ (C) or 250 nM 32 P-automodified PARP-1 with a 100-fold excess of RNA (D). Lanes: 1, core histones; 2, GST-hnRNP A1 Δ RRM2; 3, GST-hnRNP A1 UPI domain; 4, GST-hnRNP A1 C-terminal domain; 5, GST-hnRNP A2 full length; 6, GST-hnRNP A1 full length; 7, GST protein; 8, molecular-mass standards (molecular masses are given in kDa on the extreme right).

hnRNP A1 mutants did not allow us to point out single critical amino acid positions within the pADPr-binding site.

Characterization of pADPr binding on hnRNP conserved domains

Recombinant purified hnRNP-GST hybrid proteins were used in a pADPr-binding assay in order to characterize further the binding specificity of hnRNPs (Figure 8). A SYPRO Ruby Red blot stain of purified hnRNP conserved domains is shown in Figure 8(A). Recombinant hnRNPs constructs migrated at their expected molecular mass, but some proteolysis is apparent. As expected, GST constructs of hnRNP A1 and hnRNP A2 strongly bound pADPr (Figure 8B). Core histones and purified GST were used as positive and negative control samples respectively. Indeed, histones are strong binders of pADPr in this system, while no binding is observed for the GST protein. Specific domains of hnRNP A1 were also expressed as a hybrid protein with GST. No binding is observed for either part of the UPI or CT domains of hnRNP A1. However, a GST construct exhibiting hnRNP A1 protein with a deletion of the RRM2 does bind to pADPr. This result reveals that RRM2 is not involved in pADPr binding, because the complete deletion of this RRM did not decrease the apparent affinity. However, the UPI and CT domains constructs failed to bind pADPr. Such a result was expected with the glycine-rich CT domain, as it does not contain the putative pADPr-binding site. Surprisingly, the UPI domain construct that contains the two intact RRMs and the putative pADPr-binding site did not bind any pADPr. Apparently, the glycine-rich CT domain could be required for adequate position and availability of the pADPr-binding site within the UPI domain, since pADPr-binding is restored when hnRNP constructs containing both the CT domain and pADPr-binding site are present [full-length hnRNP A1 and hnRNP A1 Δ RRM2]. Unfortunately, structural information is not available for the intact hnRNP A1 protein, and therefore the possible structural effect of the CT domain on pADPr-binding is not known.

The unchanged pADPr-binding affinity of the hnRNP A1 lacking a RRM2 sequence is interesting, because it segregates RNA binding from pADPr binding. This observation is also confirmed by competition assays using excess RNA or NAD $^{+}$ over pADPr (Figures 8B and 8C). Exactly the same binding pattern and intensity was observed, indicating that there is no change in pADPr-binding affinity. These interactions are very specific, because they are observed under physiological or high-stringency conditions (1 M NaCl). Moreover, the inability of competitive RNA or NAD $^{+}$ to affect the binding efficiency is also a strong indication that pADPr binding to hnRNPs is a specific process.

DISCUSSION

Using a proteomic approach, we have identified a new family of proteins that bind to pADPr, namely hnRNPs. Prasad et al. [33] have previously reported the ADP-ribosylation of hnRNP A1 and hnRNP A2/B1. At that time, it was impossible to determine whether these hnRNPs were modified by single or multiple covalently linked ADP-ribose units. In view of our results it is clear, however, that hnRNPs can establish a strong non-covalent link with pADPr. We demonstrate for the first time that hnRNPs have specific affinity to pADPr resulting from a non-covalent interaction mediated by a consensus motif within the amino acids sequence. Although the function of non-covalent interactions between free pADPr and proteins is poorly understood, it has been demonstrated that non-covalent interactions between the pADPr nuclear proteins such as p53 and histones are very stable and could modify the functional properties of these and other proteins in living cells [34–36]. Therefore, these previous results substantiate our finding that hnRNPs are specific non-covalent pADPr binders.

The interaction between hnRNPs and pADPr could play several cellular functional roles. Given the fact that poly(ADP-

ribosyl)ation is an early response to DNA damage, the affinity of hnRNPs to pADPr, which plays important roles in RNA maturation and translocation, is intriguing. For instance, the interaction of hnRNPs with pADPr could be a mechanism for controlling RNA maturation, because hnRNPs are involved in RNA splicing [8].

PARP-1 translocation to the nucleoplasm during DNA damage and apoptosis [37,38] could affect RNA maturation through the interaction of automodified PARP-1 with hnRNPs. The interaction proposed with PARP-1 is consistent with the nuclear distribution of hnRNPs. The previous finding that hnRNP A1 is specifically cleaved by caspase-3 proteases also supports the link between hnRNPs and apoptosis events [19].

PARP-1 has been shown to be involved in DNA damage signalling through its interaction with p53, p21 and other proteins [39]. Indeed, PARP-1^{-/-} cells are extremely sensitive to γ -irradiation [40]. Part of this signalling mechanism could be the retention of hnRNPs in the nucleus during the peak of pADPr synthesis by PARP-1 in response to DNA-damaging agents. Thus, pADPr synthesis could affect cytoplasmic shuttling during apoptosis and DNA damage. Also, because there is a very large increase of pADPr levels during apoptosis, hnRNP packaging and transport out of the nucleus could be affected by pADPr binding.

We hope that our results will provide important insights into the potential modulation role that pADPr binding to hnRNPs may play in mRNA processing and trafficking into the cell, especially during apoptosis-related cellular events.

This work was supported by the Canadian Institute of Health Research. We thank Isabelle Kelly for peptide synthesis. We also thank Dr Eric Winstall for informative discussions.

REFERENCES

- Juarez-Salinas, H., Sims, J. L. and Jacobson, M. K. (1979) Poly(ADP-ribose) levels in carcinogen-treated cells. *Nature (London)* **282**, 740–741.
- Lindahl, T., Satoh, M. S., Poirier, G. G. and Klungland, A. (1995) Post-translational modification of poly(ADP-ribose) polymerase induced by DNA strand breaks. *Trends Biochem. Sci.* **20**, 405–411.
- D'Amours, D., Desnoyers, S., D'Silva, I. and Poirier, G. G. (1999) Poly(ADP-ribose)ylation reactions in the regulation of nuclear functions. *Biochem. J.* **342**, 249–268.
- Yu, S. W., Wang, H., Poitras, M. F., Coombs, C., Bowers, W. J., Federoff, H. J., Poirier, G. G., Dawson, T. M. and Dawson, V. L. (2002) Mediation of poly(ADP-ribose) polymerase-1-dependent cell death by apoptosis-inducing factor. *Science* **297**, 259–263.
- Panda, S., Poirier, G. G. and Kay, S. A. (2002) *tej* defines a role for poly(ADP-ribose)ylation in establishing period length of the *Arabidopsis* circadian oscillator. *Dev. Cell* **3**, 51–61.
- Althaus, F. R. and Richter, C. (1987) ADP-Ribosylation of Proteins: Enzymology and Biological Significance, Springer-Verlag, Berlin.
- Ogata, N., Ueda, K., Kawaichi, M. and Hayaishi, O. (1981) Poly(ADP-ribose) synthetase, a main acceptor of poly(ADP-ribose) in isolated nuclei. *J. Biol. Chem.* **256**, 4135–4137.
- Krecic, A. M. and Swanson, M. S. (1999) hnRNP complexes: composition, structure, and function. *Curr. Opin. Cell Biol.* **11**, 363–371.
- Bennett, M., Pinol-Roma, S., Staknis, D., Dreyfuss, G. and Reed, R. (1992) Differential binding of heterogeneous nuclear ribonucleoproteins to mRNA precursors prior to spliceosome assembly *in vitro*. *Mol. Cell. Biol.* **12**, 3165–3175.
- Guillonnet, F., Guieysse, A. L., Le Caer, J. P., Rossier, J. and Praseuth, D. (2001) Selection and identification of proteins bound to DNA triple-helical structures by combination of 2D-electrophoresis and MALDI-TOF mass spectrometry. *Nucleic Acids Res.* **29**, 2427–2436.
- Hay, D. C., Kemp, G. D., Dargemont, C. and Hay, R. T. (2001) Interaction between hnRNP A1 and I κ B α is required for maximal activation of NF- κ B-dependent transcription. *Mol. Cell. Biol.* **21**, 3482–3490.
- Michael, W. M., Siomi, H., Choi, M., Pinol-Roma, S., Nakielnny, S., Liu, Q. and Dreyfuss, G. (1995) Signal sequences that target nuclear import and nuclear export of pre-mRNA-binding proteins. *Cold Spring Harbor Symp. Quant. Biol.* **60**, 663–668.
- Michael, W. M., Choi, M. and Dreyfuss, G. (1995) A nuclear export signal in hnRNP A1: a signal-mediated, temperature-dependent nuclear protein export pathway. *Cell* **83**, 415–422.
- Pinol-Roma, S. and Dreyfuss, G. (1992) Shuttling of pre-mRNA binding proteins between nucleus and cytoplasm. *Nature (London)* **355**, 730–732.
- Ford, L. P., Wright, W. E. and Shay, J. W. (2002) A model for heterogeneous nuclear ribonucleoproteins in telomere and telomerase regulation. *Oncogene* **21**, 580–583.
- Izaurralde, E., Jarmolowski, A., Beisel, C., Mattaj, J. W., Dreyfuss, G. and Fischer, U. A. (1997) Role for the M9 transport signal of hnRNP A1 in mRNA nuclear export. *J. Cell Biol.* **137**, 27–35.
- Casas-Finet, J. R., Smith, Jr, J. D., Kumar, A., Kim, J. G., Wilson, S. H. and Karpel, R. L. (1993) Mammalian heterogeneous ribonucleoprotein A1 and its constituent domains. Nucleic acid interaction, structural stability and self-association. *J. Mol. Biol.* **229**, 873–889.
- Mayeda, A., Munroe, S. H., Caceres, J. F. and Krainer, A. R. (1994) Function of conserved domains of hnRNP A1 and other hnRNP A/B proteins. *EMBO J.* **13**, 5483–5495.
- Brockstedt, E., Rickers, A., Kostka, S., Laubersheimer, A., Dörken, B., Wittmann-Liebold, B., Bommer, K. and Otto, A. (1998) Identification of apoptosis-associated proteins in a human Burkitt lymphoma cell line. Cleavage of heterogeneous nuclear ribonucleoprotein A1 by caspase 3. *J. Biol. Chem.* **273**, 28057–28064.
- Thiede, B., Dimmler, C., Siejak, F. and Rudel, T. (2001) Predominant identification of RNA-binding proteins in Fas-induced apoptosis by proteome analysis. *J. Biol. Chem.* **276**, 26044–26050.
- Thiede, B., Siejak, F., Dimmler, C. and Rudel, T. (2002) Prediction of translocation and cleavage of heterogeneous ribonuclear proteins and Rho guanine nucleotide dissociation inhibitor 2 during apoptosis by subcellular proteome analysis. *Proteomics* **2**, 996–1006.
- Mortz, E., Krogh, T. N., Vorum, H. and Gorg, A. (2001) Improved silver staining protocols for high sensitivity protein identification using matrix-assisted laser desorption/ionization-time of flight analysis. *Proteomics* **1**, 1359–1363.
- Ménard, L. and Poirier, G. G. (1987) Rapid assay of poly(ADP-ribose) glycohydrolase. *Biochem. Cell. Biol.* **65**, 668–673.
- Zahradka, P. and Ebisuzaki, K. (1984) Poly(ADP-ribose) polymerase is a zinc metalloenzyme. *Eur. J. Biochem.* **142**, 503–509.
- Brochu, G., Duchaine, C., Thibeault, L., Laguerre, J., Shah, G. M. and Poirier, G. G. (1994) Mode of action of poly(ADP-ribose) glycohydrolase. *Biochim. Biophys. Acta* **1219**, 342–350.
- Shah, G. M., Poirier, D., Duchaine, C., Brochu, G., Desnoyers, S., Laguerre, J., Verreault, A., Hotlack, J. C., Kirkland, J. B. and Poirier, G. G. (1995) Methods for biochemical study of poly(ADP-ribose) metabolism *in vitro* and *in vivo*. *Anal. Biochem.* **227**, 1–13.
- Pleschke, J. M., Kleczkowska, H. E., Strohm, M. and Althaus, F. R. (2000) Poly(ADP-ribose) binds to specific domains in DNA damage checkpoint proteins. *J. Biol. Chem.* **275**, 40974–40980.
- Panzeter, P. L., Zweifel, B., Malanga, M., Waser, S. H., Richard, M. and Althaus, F. R. (1993) Targeting of histone tails by poly(ADP-ribose). *J. Biol. Chem.* **268**, 17662–17664.
- Affar, E. B., Duriez, P. J., Shah, R. G., Winstall, E., Germain, M., Boucher, C., Bourassa, S., Kirkland, J. B. and Poirier, G. G. (1999) Immunological determination and size characterization of poly(ADP-ribose) synthesized *in vitro* and *in vivo*. *Biochim. Biophys. Acta* **1428**, 137–146.
- Ben-David, Y., Bani, M. R., Chabot, B., De Koven, A. and Bernstein, A. (1992) Retroviral insertions downstream of the heterogeneous nuclear ribonucleoprotein A1 gene in erythroleukemia cells: evidence that A1 is not essential for cell growth. *Mol. Cell. Biol.* **12**, 4449–4455.
- LaBranche, H., Dupuis, S., Ben-David, Y., Bani, M. R., Wellinger, R. J. and Chabot, B. (1998) Telomere elongation by hnRNP A1 and a derivative that interacts with telomeric repeats and telomerase. *Nat. Genet.* **19**, 199–202.
- Combet, C., Blanchet, C., Geourjon, C. and Deleage, G. (2000) NPS@: network protein sequence analysis. *Trends Biochem. Sci.* **25**, 147–150.
- Prasad, S., Walent, J. and Dritschilo, A. (1994) ADP-ribosylation of heterogeneous ribonucleoproteins in HeLa cells. *Biochem. Biophys. Res. Commun.* **204**, 772–779.
- Panzeter, P. L., Realini, C. A. and Althaus, F. R. (1992) Noncovalent interactions of poly(adenosine diphosphate ribose) with histones. *Biochemistry* **31**, 1379–1385.
- Nozaki, T., Masutani, M., Akagawa, T., Sugimura, T. and Esumi, H. (1994) Non-covalent interaction between poly(ADP-ribose) and cellular proteins: an application of a poly(ADP-ribose)-Western blotting method to detect poly(ADP-ribose) binding on protein-blotted filter. *Biochem. Biophys. Res. Commun.* **198**, 45–51.

- 36 Malanga, M., Pleschke, J. M., Kleczkowska, H. E. and Althaus, F. R. (1998) Poly(ADP-ribose) binds to specific domains of p53 and alters its DNA binding functions. *J. Biol. Chem.* **273**, 11839–11843
- 37 Alvarez-Gonzalez, R., Spring, H., Muller, M. and Burkle, A. (1999) Selective loss of poly(ADP-ribose) and the 85-kDa fragment of poly(ADP-ribose) polymerase in nucleoli during alkylation-induced apoptosis of HeLa cells. *J. Biol. Chem.* **274**, 32122–32126
- 38 Germain, M., Affar, E. B., D'Amours, D., Dixit, V. M., Salvesen, G. S. and Poirier, G. G. (1999) Cleavage of automodified poly(ADP-ribose) polymerase during apoptosis. Evidence for involvement of caspase-7. *J. Biol. Chem.* **274**, 28379–28384
- 39 Wesierska-Gadek, J. and Schmid, G. (2001) Poly(ADP-ribose) polymerase-1 regulates the stability of the wild-type p53 protein. *Cell. Mol. Biol. Lett.* **6**, 117–140
- 40 Ménissier de Murcia, J., Niedergang, C., Trucco, C., Ricoul, M., Dutrillaux, B., Mark, M., Oliver, F. J., Masson, M., Dierich, A., LeMeur, M. et al. (1997) Requirement of poly(ADP-ribose) polymerase in recovery from DNA damage in mice and in cells. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 7303–7307
- 41 Vitali, J., Ding, J., Jiang, J., Zhang, Y., Krainer, A. R. and Xu, R. M. (2002) Correlated alternative side chain conformations in the RNA-recognition motif of heterogeneous nuclear ribonucleoprotein A1. *Nucleic Acids Res.* **30**, 1531–1538

Received 25 October 2002/17 December 2002; accepted 8 January 2003

Published as BJ Immediate Publication 8 January 2003. DOI 10.1042/BJ20021675